



计算机科学

COMPUTER SCIENCE

基于改进YOLOv4的磁瓦缺陷检测算法

张晓晓, 邓承志, 吴朝明, 曹春阳, 胡诚

引用本文

张晓晓, 邓承志, 吴朝明, 曹春阳, 胡诚. [基于改进YOLOv4的磁瓦缺陷检测算法](#)[J]. 计算机科学, 2023, 50(11A): 230100100-7.

ZHANG Xiaoxiao, DENG Chengzhi, WU Zhaoming, CAO Chunyang, HU Cheng. [Magnetic Tile Defect Detection Algorithm Based on Improved YOLOv4](#) [J]. Computer Science, 2023, 50(11A): 230100100-7.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于机器视觉的超声相控阵缺陷检测研究](#)

Study on Ultrasonic Phased Array Defect Detection Based on Machine Vision

计算机科学, 2023, 50(11A): 230200150-6. <https://doi.org/10.11896/jsjcx.230200150>

[一种面向工业产品表面缺陷图像的色调增强方法](#)

Hue Augmentation Method for Industrial Product Surface Defect Images

计算机科学, 2023, 50(11A): 230200089-6. <https://doi.org/10.11896/jsjcx.230200089>

[改进YOLOv5s的轻量化钢材表面缺陷检测模型](#)

Improved YOLOv5s Lightweight Steel Surface Defect Detection Model

计算机科学, 2023, 50(11A): 230900113-7. <https://doi.org/10.11896/jsjcx.230900113>

[基于多尺度特征融合的遥感图像建筑物提取算法研究](#)

Study on Building Extraction Algorithm of Remote Sensing Image Based on Multi-scale Feature Fusion

计算机科学, 2023, 50(9): 202-209. <https://doi.org/10.11896/jsjcx.220800086>

[基于低开销可变形卷积的MobileNet再轻量化方法](#)

Re-lightweight Method of MobileNet Based on Low-cost Deformable Convolution

计算机科学, 2022, 49(12): 312-318. <https://doi.org/10.11896/jsjcx.211200036>

基于改进 YOLOv4 的磁瓦缺陷检测算法

张晓晓 邓承志 吴朝明 曹春阳 胡 诚

南昌工程学院信息工程学院 南昌 330099

(zhangxiaoxiao_666@163.com)

摘要 磁瓦在生产制造过程中会因为工艺问题产生各种不同的缺陷,传统检测算法检测速度慢、精度低,为了实现磁瓦表面缺陷快速有效的检测,文中提出了一种改进 YOLOv4 算法的磁瓦缺陷检测方法。首先将 scSE 注意力模块嵌入特征提取主干网络中的 CSPnet 的残差单元中,增强小目标的空间特征和通道特征;其次,采用空洞卷积空间池化金字塔(ASPP)模块代替原有 SPP 模块,增大卷积核感受野,更多地保留图像细节并增强信息相关性;最后,在颈部部分用深度可分离卷积替换 5 次卷积块中的传统卷积,以此来更好地对特征信息进行提取,减小模型的参数量。实验结果表明,改进的 YOLOv4 算法对磁瓦表面缺陷检测的平均精度值达到 96.67%,检测速度为 44 ms,模型大小为 249 MB,明显优于原始算法,具有较高的检测精度和实用性。

关键词: 缺陷检测;YOLOv4;scSE 注意力;空洞卷积池化;深度可分离

中图分类号 TP391

Magnetic Tile Defect Detection Algorithm Based on Improved YOLOv4

ZHANG Xiaoxiao, DENG Chengzhi, WU Zhaoming, CAO Chunyang and HU Cheng

School of Information Engineering, Nanchang Institute of Technology, Nanchang 330099, China

Abstract Various defects occur in the manufacturing process of magnetic tiledue to process problems, and traditional detection algorithms have slow detection speed and low accuracy. In order to achieve fast and effective detection of surface defects of magnetic tiles, this paper proposes a defect detection method for magnetic tiles with improved YOLOv4 algorithm. Firstly, the scSE attention module is embedded in the residual unit of CSPnet in the feature extraction backbone network to enhance the spatial features and channel features of small targets. Secondly, the empty convolutional space pooling pyramid(ASPP) module is used instead of the original SPP module to increase the perceptual field of convolutional kernel, retain more image details and enhance information relevance. Finally, the traditional convolution in the five convolution blocks is replaced by the depth-separable convolution in the neck part to better extract the feature information and reduce the number of parameters of the model. Experimental results show that the improved YOLOv4 algorithm achieves an average accuracy value of 96.67%, a detection speed of 44 ms, and a model size of 249 MB. It is significantly better than the original algorithm and has higher detection accuracy and practicality.

Keywords Defect detection, YOLOv4, scSE attention, Void convolution pooling, Depth-separable

1 引言

随着人们对汽车性能及使用感受要求的不断增加,电子化配置在汽车上的应用越来越丰富,因此电机在汽车上的应用明显增加,发动机中电机的重要零部件磁瓦表面的缺陷会影响发动机的使用寿命。为了保证电机的可靠运行,电机的主要部件磁瓦的性能要求就有所提升。磁瓦缺陷类型种类不一,本文主要针对 5 种类型的磁瓦缺陷,包括缺陷较大的有磨损、崩缺、不规则类型,缺陷区域较细小的有孔洞、裂纹类型,其中裂纹深浅不一,较小的线宽一般为 0.2mm~1mm,长度不固定。传统的检测方法多依赖人工目检,这种检测方法成本高、效率低,受检测者主观因素影响标准不统一,且易损害检测人员的身体健康。因此,高效的磁瓦表面缺陷检测技术成为了制约磁瓦行业发展亟待解决的关键问题^[1]。

现有的基于 CNN 缺陷检测网络算法在结构上可分为

两个阶段:单阶段和两阶段。为了提高现有检测网络的精确度, GIRSHICK 等先后提出了 R-CNN(2014)^[2], Fast R-CNN(2015)^[3] 和 Faster R-CNN(2017)^[4] 等检测器,这一系列检测器从结构上可划分为两阶段算法,由于这类检测器在提高精度的同时带来了速度慢的问题,因此不适用于实时监测需求较高的磁瓦缺陷检测。

为了提高检测器的效率,单阶段目标检测算法应运而生。此种检测器能直接预测输入图片的位置信息与类别概率,可以兼顾检测精度和检测速度,满足实时性要求较高的检测任务。YOLO 算法属于单阶段算法的一种,其由 Joseph 等于 2016 年首次提出^[5],后续作者对其进行不断改进,逐渐衍生出 YOLOv2^[6] 与 YOLOv3,检测性能得到稳步提升,引起学术界的广泛关注与研究。Cao 等在 YOLOv3 的主干网络和特征融合网络中加入 SPP 模块^[7],并使用 Mish 作为激活函数,解决了因局部信息缺失导致小物体检测性能差的问题,在提

基金项目:江西级研究生创新专项基金项目(YC2021-S184);南昌工程学院研究生创新专项计划项目(YJSCX202130)

This work was supported by the Jiangxi Province Postgraduate Innovation Special Fund Project(YC2021-S184) and Nanchang Institute of Technology Postgraduate Innovation Special Fund Project(YJSCX202130).

通信作者:邓承志(dengcz@nit.edu.cn)

高检测速度的同时将 mAP 提升了 13 个百分点。2020 年 Alexey 等在传统 YOLO 的基础上极具创新性地融合了多个实用模块,提出了 YOLOv4 算法^[8],与以往的 YOLO 算法相比检测效率和检测效果都得到巨大的提升。在此基础上,Xie 等通过引入多级残差混合注意力模块(MRHAM)到 YOLOv4 的网络结构上,并通过 K-means++ 聚类算法^[9]聚类先验框,从而提高模型对小目标的检测性能,再结合迁移学习和数据增强等方法增强模型的泛化能力^[10],这种改进的算法在测试集上的 mAP 达到 99.71%,相比原 YOLOv4 提升了 3.50%。

上述目标检测算法在提高模型识别准确率和检测效率方面都取得了不错的成果。然而,对于对比度低、尺寸变化大的微小目标仍然存在漏检、错检等问题。针对以上算法的不足,本文对其进行了改进,主要贡献如下:

- 1) 将 scSE 注意力机制嵌入在 CSPDarknet53 网络的残差块,能增强小目标的空间特征和通道特征,解决微小磁瓦缺陷特征损失严重的问题,提升了算法对小目标的检测性能。
- 2) 为解决 SPP 结构伴随图像分辨率损失的问题,采用 ASPP(空洞卷积空间池化金字塔)模块代替原有 SPP 模块,在增大卷积核感受野的同时更多地保留图像细节并增强信息相关性,微小缺陷的漏检、误检问题得到明显改善。
- 3) 在 Neck 部分引入深度可分离卷积,减少模型的参数量,以提高磁瓦表面缺陷检测的速度。

2 YOLOv4 算法介绍

YOLO 作为 One-Stage 目标检测算法的一种^[11],于 2016 年由 Redmon 等提出,是目前应用最广泛的表面缺陷检测模型之一,已发展为多个版本。

YOLOv4 是 YOLOv3 的升级版,实现了检测速度和准确

率的最优平衡。YOLOv4 网路由 CSPDarknet53, SPP, PANet, YOLOv3 头部组成,输入的是 416×416 固定大小的图像,并将输入图像分割为 $s \times s$ 个网格,该网格的每个单元格负责预测固定数量的带有置信度的边界框,用于实现目标的分类与定位,从而实现端到端的目标检测。

YOLOv4 在输入端引入了 Mosaic 数据增强^[12],在训练时取出 4 张图片通过随机排列、裁剪以及缩放等形式进行拼接,丰富数据集的同时提升了对小目标的检测能力。YOLOv4 的主干网络是 CSPDarknet-53^[13],是一种性能更好的残差结构。在 Darknet-53 网络的基础上融合了 CSPNet 结构^[14],对各个部分进行了改进,总共包含 5 个 CSP 模块,每个模块使用 3×3 的卷积核、步长为 2 进行下采样操作,最终将输入图像为 608×608 大小的特征图变为 19×19 的特征图,实现了图像的特征提取,在减小计算量的同时保持了准确率。

YOLOv4 在 Neck 结构中不仅采用了空间金字塔池化(SPP)^[15],还引入了路径聚合网络(PANet)^[16]模块。SPP 模块首先使用 4 个不同大小的核进行最大池化,再将 4 个尺度的特征层进行拼接,使得图像特征能够充分融合,扩大网络感受野,获取重要区域的信息。YOLOv3 中使用特征金字塔网络(Feature Pyramid Network, FPN)^[17],通过自上而下的方式将高层信息通过上采样的方式进行特征融合,为了进一步提高特征提取能力,YOLOv4 在此基础上添加了 PAN 结构, PANet 部分采用 FPN+PAN 的方式对特征图进行多尺度融合以及反复特征变换,两者结合的操作加速了不同尺度的特征融合,进一步提升了网络表征能力。

YOLO-Head 检测头依然采用 YOLOv3 的多尺度思想进行预测,接收到从 PANet 输出的 3 个不同尺度的特征层,对其特征进行预测,输出检测结果。

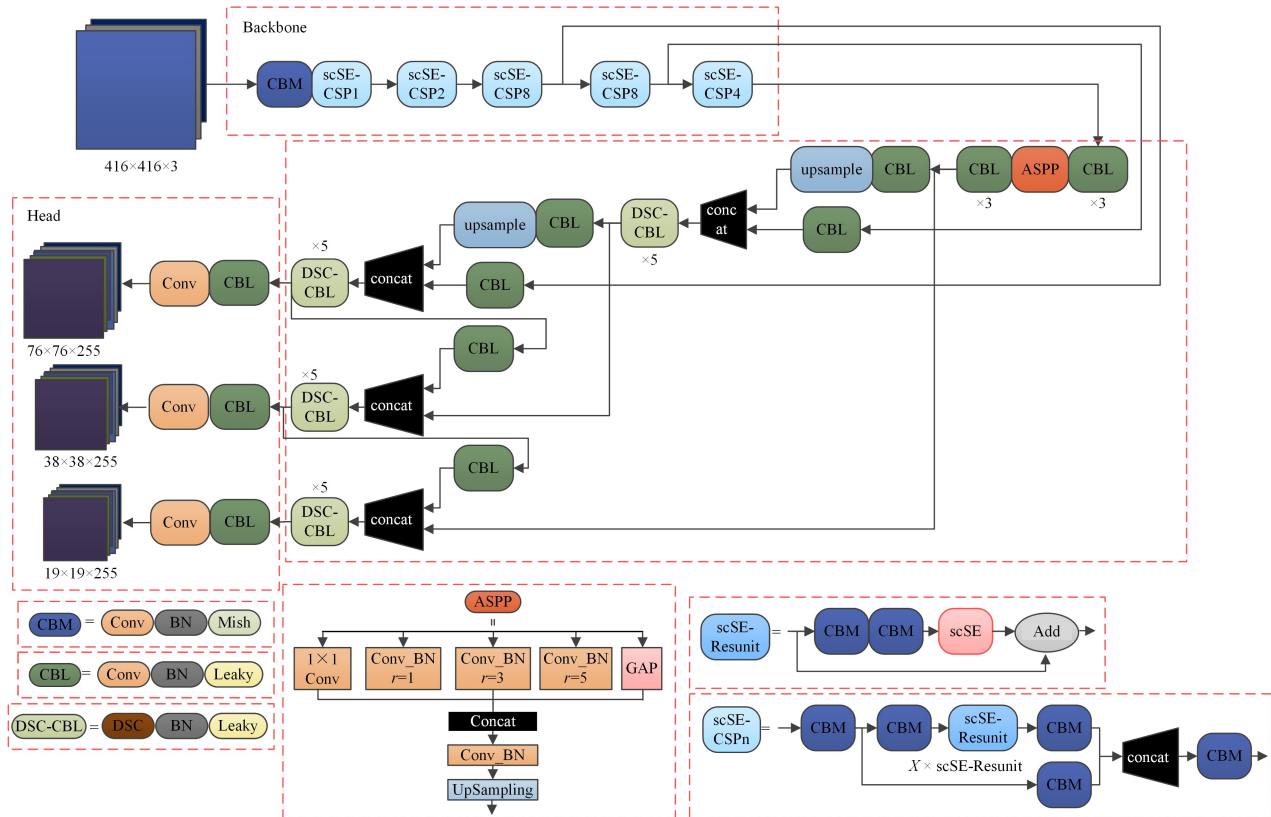


图 1 改进的 YOLOv4 模型结构

Fig. 1 Improved YOLOv4 model structure

3 改进的 YOLOv4 模型

YOLOv4 提出了很多改进策略,然而,由于单阶段网络集成了缺陷的识别与定位,在大大提升检测速度的同时,对于微小、几何形态多变的缺陷无法兼顾检测精度,并且对于对比度低、尺寸变化大的微小目标仍然存在漏检、错检的问题。

针对上述问题,本文提出了改进 YOLOv4 算法,用于磁瓦缺陷检测。首先,本文通过增加注意力机制 scSE 模块来提升网络的特征提取能力。其次,采用空洞卷积空间池化金字塔模块代替原有 SPP 模块,增强网络整体的感受野,构建有利于小目标缺陷识别的特征融合方式。最后,使用深度可分离卷积替代传统的 3×3 卷积核,有效地减小了参数数量,提高了网络的检测速度。从以上 3 个方面对 YOLOv4 算法进行改进,整体提高模型的检测精度及速度。图 1 给出了改进的 YOLOv4 模型结构。

3.1 scSE-CSPnet 模块

注意力机制模拟人类大脑信号处理机制,在扫描全局图像时,利用注意力机制使网络关注重要的目标区域,抑制不相关区域^[18]。

cSENet 为通道注意力机制,模型结构由压缩、激励、注意 3 部分组成,通过为每个通道赋予不同的权重,使得网络关注图片的重要区域,以此提高网络的表现力。其操作步骤为:首先,将输入特征图进行全局平均池化操作,对其通道压缩输出一个尺寸为 $1 \times 1 \times C$ 的向量;其次,将向量输入到两个全连接层与 Relu 激活层,此时完成对通道依赖项的编码;最后,通过 Sigmoid 激活函数^[19]使其值介于 $0 \sim 1$ 之间,并用于重新校准通道内的值。

sSENet 是空间注意力机制,通过学习所有通道的整体分

布,舍弃奇异的通道,得到一个权重值,根据权重赋予特征图不同的注意力。具体操作为:将输入的特征图采用 1×1 卷积核进行通道压缩并进行空间激发,然后经过 Sigmoid 激活函数将数据归一到 $0 \sim 1$ 之间,用于在空间上重新校准。

scSENet 由 cSENet 和 sSENet 组合而成,并行空间和通道模块分别沿空间和通道重新校准特征映射,然后结合输出。激发特征图在空间和通道上获得更重要信息,这种重新校准鼓励网络学习更有意义的特征图。图 2 给出了 scSE 模块的结构。

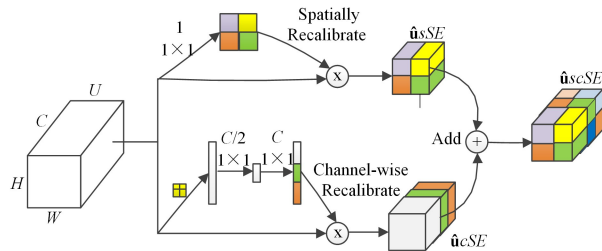


图 2 scSE 模块结构

Fig. 2 scSE module structure

CSP-Darknet53 是局部跨阶段残差网络,每个 CSPnet 都是通过融合有限感受野范围内获取的空间特征和通道特征,来构建相应的特征向量,但并未区分各个特征通道的重要性。scSE 注意力机制可以解决卷积池化过程中由通道重要性差异所引起的信息损失问题。因此,本文在 YOLOv4 的骨干网嵌入 scSE 注意力模块,将 scSE 注意力模块嵌入 CSPnet 的残差单元,先经过残差操作深入提取特征,再将特征图通过注意力机制突出关键特征,弱化无关信息,最后合并输出。改进后的 scSE-CSPnet 模块结构如图 3 所示。

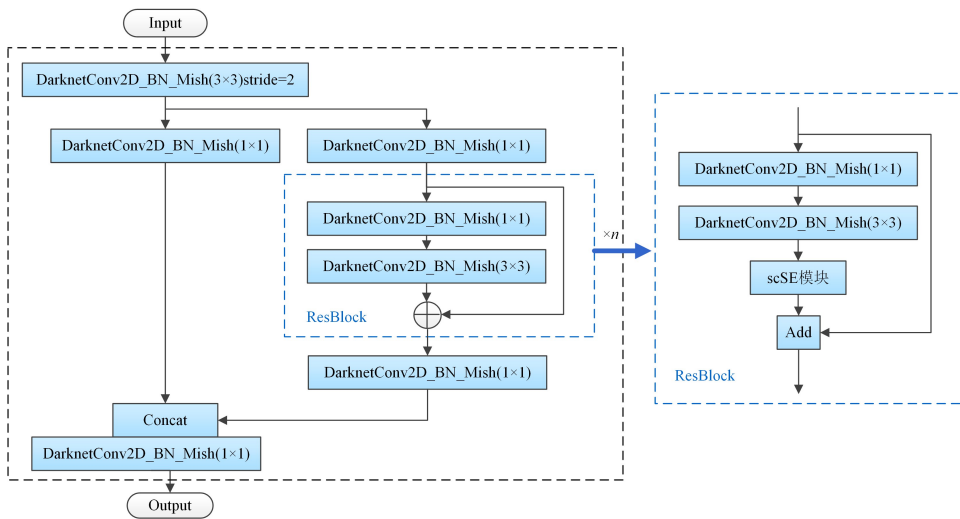


图 3 scSE-CSPnet 模块结构

Fig. 3 scSE-CSPnet module structure

3.2 空洞卷积空间池化金字塔

通常神经网络会通过增加卷积层和池化层来实现扩大感受野^[20],缺点是会使网络深度增加,甚至可能出现梯度消失的问题,不利于特征的传递。更严重的是,特征图的尺寸会随着池化层的增多而越来越小,越容易丢失特征信息,使后面的操作未能很好地发挥作用,导致网络精度不能明显提高。

ASPP 是将 SPP 中的池化替换成空洞卷积操作,与全局平均池化并联,可以扩大感受野,从而获取多尺度信息。通过

在卷积层中引入扩张率,表示卷积核间零值的数量,在特征提取时实现跨像素提取。因此,本文引入 ASPP 模块代替原有的 SPP 模块,以获取更完整的局部特征,提高对微小缺陷的磁瓦检测能力。空洞卷积池化金字塔如图 4 所示。

为了提取多尺度信息,获取不同尺度的感受野,ASPP 模块使用不同的填充以及膨胀,而卷积时的填充及膨胀是由扩张率控制的。扩张率对卷积层的感受野的影响如式(1)所示:

$$R = (r' - 1) \times (k - 1) + k \quad (1)$$

其中, k 为卷积核大小, r' 为扩张率。随着感受野范围变大, 可获取的特征信息越丰富, 但对于局部特征信息的提取难度也会增加。自适应均值池化层对图像进行全局特征信息提取, 再经 1×1 卷积后上采样恢复至输入特征图像大小。ASPP 模块将不同尺度的特征信息与全局特征串接起来, 利用卷积来改变输出的通道数, 从而得到网络的输出特征图。ASPP 模块改善了单一空洞卷积所带来的信息损失, 在增强信息相关性的同时保留了更多的图像特征, 提升了对大小目标的检测效果。

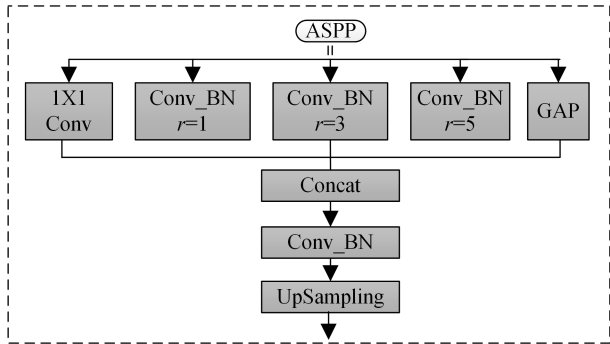


图 4 空洞卷积空间池化金字塔
Fig. 4 Atrous spatial pyramid pooling

3.3 深度可分离卷积

深度可分离卷积分为逐通道卷积和逐点卷积两个过程, 将这两部分组合进行特征提取。它是一种常规卷积的替代方式, 计算过程保证了空间信息的捕捉和跨通道信息的融合, 在基本保留卷积核的表征学习能力的同时减少了权重系数的个数, 提升了卷积核参数的使用效率。DSC 的详细结构示意图如图 5 所示。

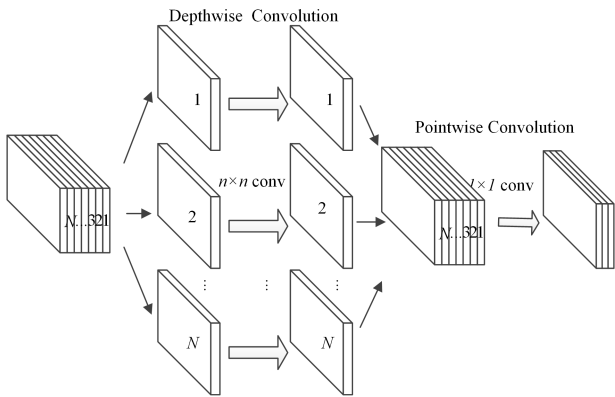


图 5 DSC 结构示意图
Fig. 5 DSC structure diagram

PANet 自底向上增强路径, 缩短了底层和顶层间的距离, 同时将高低层的特征进行融合, 从而丰富了特征信息, 提升了检测效果。以深度可分离卷积替换 5 次卷积块中的卷积核大小为 3×3 的传统卷积, 可以有效避免信息丢失, 更有效地进行特征提取, 提升缺陷的检测能力, 同时降低参数量, 提高检测速度。

4 实验

本文的实验硬件配置为: AMD Ryzen9 3900X 12 核 24 线程处理器, GeForce RTX3090 显卡, 32GB 运行内存, 500GB NVME 固态硬盘; 系统平台为 64 位 Windows11 专业版, 使用 Pycharm 集成开发环境, 配置的 CUDA 版本为 11.1, CUDNN

版本为 8.0.4.30, Pytorch 版本为 1.8.1。

4.1 实验数据集

为了检验算法的性能, 本文使用公开的磁瓦缺陷数据集 (Magnetic tile dataset) 进行模型测试^[21]。该数据集共有 450 张磁瓦表面缺陷图像, 缺陷分为孔洞、裂纹、磨损、崩缺、不规则 5 个类别, 通过旋转、水平翻转、剪裁、明暗度调整等各种数据增强方法将数据扩充到 4267 张。实验首先利用 LabelImg 软件对增强后的缺陷图像进行标注, 然后整理成 VOC 格式, 再将制作好的数据集按 7:2:1 的方式分为训练集、验证集和测试集。

4.2 评价指标

为了检验本文算法的有效性, 选取 0.5 交并比 (IOU) 阈值下的平均精度 (AP) 与平均精度均值 (mAP) 以及 FPS 作为本文方法的评价指标, 网络模型的复杂度则用模型大小 (Model Size) 与参数量 (Parameter) 来进行评价。

准确率表示正确预测的样本占有检测出目标的比例, 表达式如式 (2) 所示:

$$Precision_{rust} = \frac{TP(TruePositive)_{rust}}{TP + FN(FalsePositive)_{rust}} \quad (2)$$

召回率代表正确预测样本占有所有目标的比例, 表达式如式 (3) 所示:

$$recall = \frac{TP(TruePositive)_{rust}}{TP + FN(FalseNegative)_{rust}} \quad (3)$$

平均准确率表示 P-R 曲线与坐标轴围成的面积, 表达式如式 (4) 所示:

$$AP_{rust} = \frac{\sum Precesion}{TP + FP(FalsePositive)_{rust}} \quad (4)$$

均值平均准确率的表达式如式 (5) 所示:

$$mAP = \frac{\sum AveragePrecesion_{rust}}{n(class)} \quad (5)$$

FPS 表示模型每秒处理图片的数量, 模型的帧率越大, 处理图片的速度越快, 模型检测速度也越快。

4.3 模型训练

模型训练从零开始, Epoch 设置为 300 次, 主要分为冻结阶段和解冻阶段的训练, 其中前 50 次迭代是冻结阶段的训练, 后 250 次迭代是解冻阶段的训练。网络优化器选择 SGD 优化器, 其中初始学习率设置为 0.01, 权值衰减设置为 0.0005。训练完的损失函数 Loss 曲线如图 6 所示。

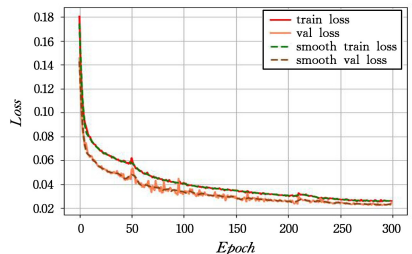


图 6 损失函数

Fig. 6 Loss function

由图 6 可知, 随着迭代的不断增加, Loss 值不断下降。在前 50 迭代中, 模型的主干处于冻结阶段时, 显存占用较小, 仅对网络进行微调, 可以加快训练效率, 也可以防止权值被破坏, 图中此阶段的 Loss 下降较快。后 250 次迭代中, 模型的主干处于解冻阶段, 占用的显存会较大, 网络参数发生变化, Loss 下降缓慢, 并逐渐趋平。最终在第 300 次迭代时, 网络收敛。

4.4 消融实验

为全面分析各个改进模块对磁瓦缺陷检测的优越性,在相同的实验环境下,本文在原 YOLOv4 的基础上设计了消融实验。实验结果如表 1 所列,其中 IOU 阈值设置为模型评估通用的 0.5,即 $mAP@0.5$ 。

表 1 消融实验结果对比分析

Table 1 Comparative analysis of ablation results

实验	scSE	ASPP	DSC	检测 mAP/%	模型 大小/MB	参数量	检测 速度/FPS
1				90.26	244	42.89×10^6	47
2	✓			94.04	256	46.52×10^6	49
3		✓		93.86	261	49.58×10^6	50
4			✓	95.31	231	38.66×10^6	45
5	✓	✓		94.52	262	53.87×10^6	52
6	✓	✓	✓	96.67	249	50.06×10^6	44

从消融实验结果可知,将 scSE 注意力模块嵌入到骨干网络的残差结构中“Add”特征融合层前能增强算法对小目标的检测能力,虽然模型大小增加了 12 MB,模型参数量增加了 3.62×10^6 ,检测速度也有所下降,但 $mAP@0.5$ 值较 YOLOv4 提升了 3.78%。由于骨干网中特征图的语义信息并不丰富,仍然保留着目标在特征图中的中、浅层的纹理信息和轮廓信息,而这些信息对于小目标的检测至关重要,在骨干网中嵌入 scSE 注意力模块虽然一定程度上牺牲了检测速度,但能更好对特征图中小目标的空间特征和通道特征进行增强,增强小目标的检测精度。

将 YOLOv4 的 SPP 模块替换成 ASPP,各个类别的 AP 都得到了一定的提升, mAP 值提升了 3.6%,模型大小增加了 17MB,模型参数量增加了 6.69×10^6 ,检测速度下降了 3FPS。该模块扩大网络的感受野,提升了对磨损、崩缺这类较大目标的检测能力,通过不同扩张率的空洞卷积,较原 YOLOv4 算法减少了特征信息的损失,保留了更多的细节特征,改善了对孔洞、裂纹这类微小目标的特征提取效果,虽然 ASPP 会使模型稍微复杂,参数增多,但能有效提高

网络的整体性能。

将传统的卷积操作替换为深度可分离卷积能有效地提高网络特征提取能力, $mAP@0.5$ 值较 YOLOv4 提升了 5.05%,模型大小减少了 13 MB,模型参数量减少了 4.23×10^6 ,检测速度提升了 2FPS,这说明与传统卷积核相比,深度可分离卷积核可以极大地降低模型的训练参数数量和计算量,能有效提升磁瓦检测速度。

实验 5 为加入 scSE 注意力模块和 ASPP 模块的实验数据。 $mAP@0.5$ 值较 YOLOv4 提升了 4.26%,模型大小增加了 18MB,模型参数量增加了 10.98×10^6 ,检测速度下降了 5FPS。可以发现:scSE 注意力模块和 ASPP 模块虽然会增加模型的复杂度,使得参数量和模型尺寸增加,检测速度也有所降低,但两者的结合能够大幅度地提升算法对磁瓦缺陷的检测精度,尤其是对于小目标缺陷检测能力的提升显著。

通过消融实验对比发现,加入各个改进模块后的实验 6 中算法性能提升最为显著,与原 YOLOv4 相比,模型大小仅增加了 5MB,模型参数量增加了 7.17×10^6 ,但 mAP 值提升了 6.41%,检测速度也提升了 3FPS。本文算法对像素占比低的微小目标的漏检误检问题有较大改善,5 类检测目标都取得了良好的检测效果,平均检测精度值提升了 6.41%,证实了本文算法对磁瓦表面微小缺陷识别的有效性。

图 7 给出了改进 YOLOv4 算法在 5 类磁瓦表面缺陷上的 AP 值。横坐标表示召回率,纵坐标表示精准率,图像中横坐标与纵坐标所围的面积表示 AP 值。可以明显看到,只有召回率与精准率都达到较高数值时,所围成的面积才会最大,AP 值才能达到最高。图 7 中对孔洞、裂纹、不规则缺陷的 AP 值可以达到 97.5% 以上,对磨损、崩缺较小的缺陷的 AP 值也可达到 90% 以上,5 类磁瓦表面缺陷的召回率均值为 96.3%。因此,从上述 AP 值和召回率可知,本节提出的算法改善了像素占比低的微小目标漏检和误检问题,提高了平均检测精度,验证了本文算法对磁瓦表面缺陷识别的有效性。

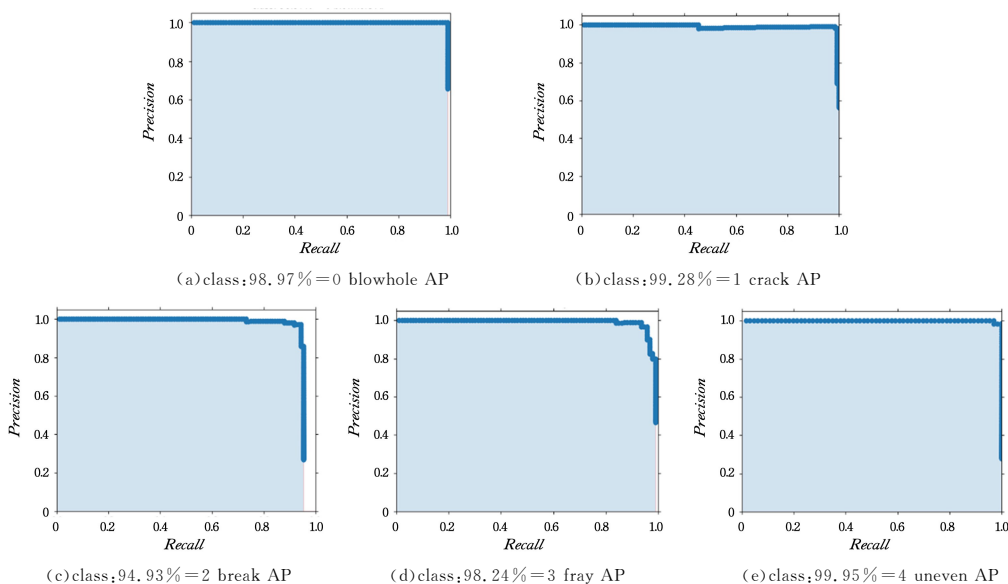


图 7 类缺陷的 AP

Fig. 7 AP for five defect types

4.5 性能对比

为进一步验证本文算法有效性,将改进后的算法与 Fas-

ter-RCNN, YOLOv3, YOLOv4, YOLOv5 算法做实验对比,在本文数据集上设置相同的超参数进行模型的训练与测试,

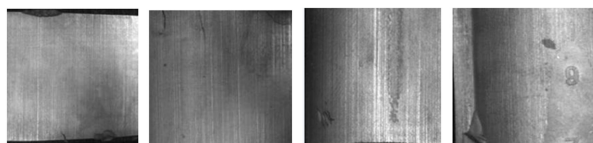
检测结果如表 2 所列。图 8 给出部分了实验的预测结果。可以发现:图 8 中第一列其他模型的算法均没有完全检测出所有破损类型缺陷,第二列 YOLOv3 模型中右上角的裂纹类型缺陷也未检测到,存在漏检的现象,而本文方法改进的 YOLOv4 算法很好地检测出了图中所有的缺陷;图 8 中第三列 Faster-RCNN 模型和 YOLOv3 模型均未检测出细小的孔

洞缺陷(见图 8(b)和图 8(c)),本文方法不仅检测出此类型缺陷,而且精度相比 YOLOv4 提升了 0.05%(见图 8(d)),说明本文算法对小目标检测有很好的效果提升;YOLOv4 模型对图 8(d)中第四列破损类别的裂缝存在错检的现象,而本文方法改进的 YOLOv4 算法能够准确地识别目标的类别(见图 8(e))。

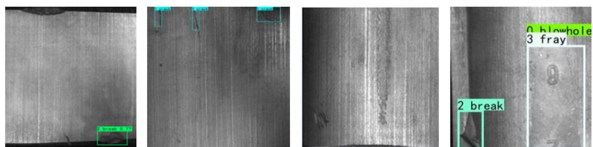
表 2 不同算法性能对比

Table 2 Performance comparison of different algorithms

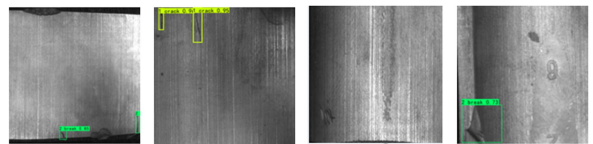
Model	Backbone	AP/%					mAP/%	检测速度/FPS
		Blowhole	crack	break	fray	uneven		
Faster-RCNN	Resnet50	68.58	73.62	79.25	96.51	90.32	81.65	68
YOLOv3	Darknet53	97.20	79.32	72.21	85.57	96.71	86.20	51
YOLOv4	CSPDarknet53	98.69	95.22	74.12	85.76	97.51	90.26	47
YOLOv5	CSPDarknet53	97.88	96.99	88.68	94.26	97.62	95.09	42
本文方法	CSPDarknet53	97.28	98.42	90.40	97.36	99.91	96.67	44



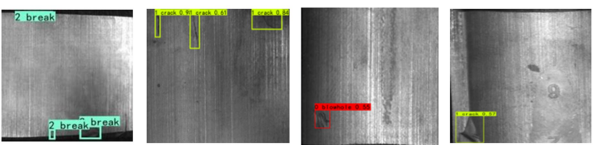
(a)原图



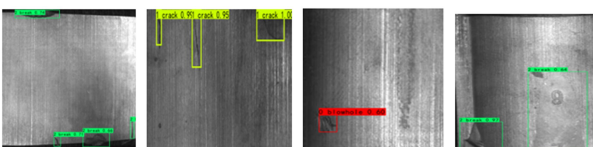
(b)Faster-RCNN 模型检测结果



(c)YOLOv3 模型检测结果



(d)YOLOv4 模型检测结果



(e)本文方法检测结果

图 8 对比实验结果

Fig. 8 Experimental results comparison

表 2 列出了本文基于 YOLOv4 提出的改进算法与 Faster-RCNN, YOLOv3, YOLOv4 和 YOLOv5 算法的性能对比。通过以上实验结果可以发现,本文方法相比主流的二阶段检测算法 FasterRCNN 和单阶段检测算法 YOLOv3, YOLOv4, YOLOv5, 平均检测精度分别提升了 15.02%, 10.47%, 6.41% 和 1.58%, 检测速度值较 FasterRCNN 和 YOLOv3, YOLOv4 算法分别降低了 24 FPS, 7 FPS, 3 FPS。mAP 值与 AP 值越大, 则检测效果越好, 精度越高, FPS 值越低, 代表单张图片的检测时间更短。实验结果表明, 本文改进方法在平均精确度 AP 和平均精度均值 mAP 方面优于 Faster-RCNN、

YOLOv3 和现有的 YOLOv4、YOLOv5 算法, 检测速度快于 FasterRCNN 和 YOLOv3、YOLOv4 算法, 略慢于 YOLOv5, 实时性较好, 基本满足了工业磁瓦检测精度及实时性需求, 有效提高了磁瓦检测的效果, 检测准确性优于其他模型, 但在实际应用中要求 99% 以上。

综合考虑检测速度、资源占比和精度等各方面的因素, 本文算法效果较优, 因此能够将其用于工业上的磁瓦缺陷检测。

结束语 针对现有网络难以快速精确识别磁瓦表面的缺陷问题, 提出了一种改进 YOLOv4 算法的磁瓦缺陷检测方法, 用于改善算法对小目标漏检、误检的问题, 从而实现对缺陷目标的准确识别。该模型将注意力机制添加到主干网络中, 用于提升网络的表征能力; 将 SPP 替换为空洞卷积空间池化金字塔, 有效地增大了网络感受野, 从而保留图像的细节特征; Neck 部分增加了深度可分离卷积, 进一步提高了检测精度。

实验结果表明, 该方法在模型尺寸、缺陷检测速度和准确性、小目标检测能力等方面均有显著提高, 相比原始 YOLOv4, 改进的 YOLOv4 算法能够较好地适用于工业下磁瓦表面的缺陷检测。

参考文献

- [1] LIU C, ZHANG J, LIN J P. Detection and Identification of Surface Defects of Magnetic Tile Based on Neural Network[J]. Surface Technology, 2019, 48(8): 330-339.
- [2] BHARATI P, PRAMANIK A. Deep learning techniques—R-CNN to mask R-CNN: a survey[C]//Computational Intelligence in Pattern Recognition(CIPR 2019). 2020:657-668.
- [3] GIRSHICK R. Fast r-cnn[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015:1440-1448.
- [4] CHEN Y, LI W, SAKARIDIS C, et al. Domain adaptive faster r-cnn for object detection in the wild[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:3339-3348.
- [5] JIANG P, ERGU D, LIU F, et al. A Review of Yolo algorithm developments[J]. Procedia Computer Science, 2022, 199: 1066-1073.
- [6] LOEY M, MANOGARAN G, TAHA M H N, et al. Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection[J]. Sus-

- tainable Cities and Society, 2021, 65: 102600.
- [7] CAO X, ZHANG F, YIC, et al. Wafer Surface Defect Detection Based On Improved YOLOv3 Network[C]// 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), 2020.
- [8] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection[J]. arXiv: 2004. 10934, 2020.
- [9] WANG P, SHI H, YANG X, et al. Three-way k-means: integrating k-means and three-way decision[J]. International Journal of Machine Learning and Cybernetics, 2019, 10: 2767-2777.
- [10] XIE H, LI Y, LI X, et al. A Method for Surface Defect Detection of Printed Circuit Board Based on Improved YOLOv4[C]// 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE). IEEE, 2021.
- [11] CHEN Q, WANG P S, CHENG A D, et al. Robust one-stage object detection with location-aware classifiers[J]. Pattern Recognition, 2020, 105.
- [12] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. Scaled-yolov4: Scaling cross stage partial network [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 13029-13038.
- [13] LI G, SUO R, ZHAO G, et al. Real-time detection of kiwifruit flower and bud simultaneously in orchard using YOLOv4 for robotic pollination[J]. Computers and Electronics in Agriculture, 2022, 193: 106641.
- [14] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020: 390-391.
- [15] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [16] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 8759-8768.
- [17] GHIASI G, LIN T Y, LE Q V. Nas-fpn: Learning scalable feature pyramid architecture for object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 7036-7045.
- [18] OBESO A M, BENOIS-PINEAU J, V? ZQUEZ M S G, et al. Visual vs internal attention mechanisms in deep neural networks for image classification and object de-tetection[J]. Pattern Recognition, 2022, 123: 108411.
- [19] KUMAR A, SINGH SODHI S. Classification of data on stacked autoencoder using modified sigmoid activation function[J]. Journal of Intelligent & Fuzzy Systems, 2023 (Preprint): 1-18.
- [20] HANANDEH S, ARDAH A, ABU-FARSAKH M. Using artificial neural network and ge-netics algorithm to estimate the resilient modulus for stabilized subgrade and propose new empirical formula[J]. Transportation Geotechnics, 2020, 24: 100358.
- [21] HUANG Y, QIU C, YUAN K. Surface defect saliency of magnetic tile[J]. The Visual Computer, 2020, 36: 85-96.



ZHANG Xiaoxiao, born in 1992, MS candidate. Her main research interests include deep learning and computer vision.



DENG Chengzhi, born in 1964, Ph. D., professor, master supervisor. His main research interests include RS- image processing and image processing.