

基于改进Deeplabv3+算法的滚珠丝杠驱动表面点蚀缺陷检测

朗朗, 陈晓琴, 刘莎, 周强

引用本文

朗朗, 陈晓琴, 刘莎, 周强. 基于改进Deeplabv3+算法的滚珠丝杠驱动表面点蚀缺陷检测[J]. 计算机科学, 2024, 51(6A): 240200058-6.

LANG Lang, CHEN Xiaoqin, LIU Sha, ZHOU Qiang. [Detection of Pitting Defects on the Surface of Ball Screw Drive Based on Improved Deeplabv3+ Algorithm](#) [J]. Computer Science, 2024, 51(6A): 240200058-6.

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)

[面向产线AI质检的少样本评测方法研究和验证](#)

Study and Verification on Few-shot Evaluation Methods for AI-based Quality Inspection in Production Lines

计算机科学, 2024, 51(6A): 230700086-8. <https://doi.org/10.11896/jsjcx.230700086>

[基于SAMNV3的滚动轴承智能故障诊断方法](#)

Intelligent Fault Diagnosis Method for Rolling Bearing Based on SAMNV3

计算机科学, 2024, 51(6A): 230700167-6. <https://doi.org/10.11896/jsjcx.230700167>

[基于LSTM和注意力机制的远程会诊需求预测](#)

Forecasting Teleconsultation Demand Based on LSTM and Attention Mechanism

计算机科学, 2024, 51(6A): 230800119-7. <https://doi.org/10.11896/jsjcx.230800119>

[融入类别标签和主题信息的用户兴趣识别方法](#)

User Interest Recognition Method Incorporating Category Labels and Topic Information

计算机科学, 2024, 51(6A): 230500169-8. <https://doi.org/10.11896/jsjcx.230500169>

[基于注意力的多尺度蒸馏异常检测](#)

Attention-based Multi-scale Distillation Anomaly Detection

计算机科学, 2024, 51(6A): 230300223-11. <https://doi.org/10.11896/jsjcx.230300223>

基于改进 Deeplabv3+ 算法的滚珠丝杠驱动表面点蚀缺陷检测

郎朗¹ 陈晓琴¹ 刘莎² 周强³

1 重庆三峡职业学院智能制造学院 重庆 404155

2 电子科技大学无线通信国家重点实验室 成都 611731

3 重庆邮电大学计算机科学与技术学院 重庆 400065

摘要 针对滚珠丝杠驱动表面背景环境复杂、点蚀缺陷目标小因而难以检测的问题,提出改进的 Deeplabv3+ 滚珠丝杠驱动表面缺陷分割算法。本算法采用 Re2Net-50 替换 Deeplabv3+ 的主干网络,显著提升了对小尺寸缺陷目标的识别能力。此外,通过在主干网络中融合特征金字塔网络 FPN,能够加强多尺度信息的提取,从而增强了对缺陷目标的精确定位。最后,本研究在 Deeplabv3+ 网络的 ASPP 模块之后引入了 Coordinate Attention 机制,能够增强模型对图像中空间和维度的关注,有效地捕获了图像中的长距离空间依赖关系。实验结果表明,与原始的 Deeplabv3+ 相比,所提算法在平均交并比 MIoU 指标上提高了 4.38%,准确率 Accuracy 提高了 5.52%,F1-score 提高了 2.74%。同时,与其他经典的语义分割算法相比,所提算法也展现出了一定的优越性。

关键词: 滚珠丝杠驱动; 缺陷检测; Deeplabv3+; 多尺度特征; 注意力机制

中图分类号 TP391

Detection of Pitting Defects on the Surface of Ball Screw Drive Based on Improved Deeplabv3+ Algorithm

LANG Lang¹, CHEN Xiaoqin¹, LIU Sha² and ZHOU Qiang³

1 School of Intelligent Manufacturing, Chongqing Three Gorges Vocational College, Chongqing 404155, China

2 The National Key Laboratory of Wireless Communications(NKLWC), University of Electronic Science and Technology, Chengdu 611731, China

3 School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract Aiming at the problems of complex background environments, small pitting defect targets, and difficulty in detection on the surface of ball screw drives, an improved Deeplabv3+ algorithm for segmenting surface defects of ball screw drives is proposed. This algorithm adopts Re2Net-50 to replace the backbone network of Deeplabv3+, significantly enhances the ability to recognize small-sized defect targets. Additionally, by integrating feature pyramid networks(FPN) into the backbone network, the algorithm effectively extracts multi-scale information, thereby improving the precise localization of defect targets. Finally, the coordinate attention mechanism is introduced after the ASPP module of the Deeplabv3+ network, enhancing the model's focus on spatial dimensions within the image and effectively capturing long-range spatial dependencies. Experimental results demonstrate that, compared to the original Deeplabv3+, the proposed algorithm shows a 4.38% improvement in the mean intersection over union(MIoU) metric, a 5.52% increase in accuracy, and a 2.74% rise in F1-score. Furthermore, when compared with other classic semantic segmentation algorithms, the proposed algorithm also exhibits certain superiority.

Keywords Ball screw drive, Defect detection, Deeplabv3+, Multi-scale features, Attention mechanism

1 引言

滚珠丝杠驱动(Ball Screw Drives, BSD)是机床中至关重要的组成部分之一,其主要功能是将旋转运动转换为直线运动,以驱动机床轴等部件。BSD 作为机床中最易磨损的零件之一^[1],也是导致机床非计划性故障的主要原因。特别是点蚀问题,它对滚珠螺杆的性能和可靠性产生了严重的负面影响。点蚀通常表现为表面上局部小范围的磨损或蚀痕,点蚀缺陷的尺寸大小从几微米到几百微米不等,从整个滚珠丝杠

表面面积来看,点蚀缺陷可能只占很小的比例。点蚀不仅会增加摩擦和磨损,降低精度,还可能导致噪音水平上升。在严重情况下,甚至可能导致滚珠螺杆完全失效。因此,对滚珠丝杠驱动表面的点蚀缺陷进行及时且准确的检测,对于保证机床等设备的正常运行至关重要。

近年来,随着深度学习技术的快速进步,基于卷积神经网络的表面缺陷检测算法也取得了显著的发展,出现了如 FCN^[2], U-Net^[3], PSPNet^[4] 等经典分割算法。许多学者在表面缺陷检测领域基于这些算法取得了卓越成果,如道路表面

基金项目:重庆市教育委员会科技研究项目(KJQN202103509);重庆市教学改革研究项目(GZ223108, GZ223113)

This work was supported by the Science and Technology Research Project of Chongqing Education Commission(KJQN202103509) and Chongqing Teaching Reform Research Project(GZ223108, GZ223113).

通信作者:郎朗(2008190232@cqszy.edu.cn)

缺陷^[5]、工业表面缺陷检测^[6]等。针对 BSD 表面缺陷的高效自动检测迫在眉睫,但此项研究一直受制于相关数据的缺乏。Schlagenhauf 等^[7]收集了一个 BSD 表面点蚀缺陷的数据集,该数据集中的缺陷目标背景复杂,缺陷像素相对较小,给准确的缺陷识别带来了较大的挑战。传统的图像处理方法在精度上难以满足工业需求,采用基于深度学习的经典分割网络能够有效检测出缺陷区域,但是针对较小的点蚀缺陷的识别上的能力仍需进一步提升。基于此,本文提出了一种改进的 Deeplabv3+^[8]滚珠丝杠驱动表面点蚀缺陷分割算法。通过替换 Deeplabv3+编码器阶段的主干网络,采用了具有更丰富多尺度信息提取能力的 Re2Net 结构^[9]进行缺陷特征提取。同时,通过在 Deeplabv3+编码器阶段深层特征阶段采用 FPN(Feature Pyramid Networks)结构^[10]进行特征融合,显著提升了对小尺寸缺陷的识别能力。此外,本研究在 ASPP 模块之后添加了卷积注意力机制 Coordinate Attention^[11],既保

留了局部特征的细致度,又有效捕捉了长距离的空间依赖关系,从而更准确地识别和定位图像中的缺陷区域。通过与 DeepLabv3+, U-Net, PSPnet, fastSCNN^[12], SwinTransformer^[13], Segformer^[14]等经典算法及本文的改进算法进行比较,证明了本文算法在 BSD 点蚀缺陷检测上的准确性。

2 DeepLabv3+ 算法结构

Deeplabv3+算法^[8]是由 Google 的研究团队开发的深度学习模型,主要用于图像的语义分割。其网络结构如图 1 所示。(1)Deeplabv3+利用空洞卷积来扩大感受野,使模型能够捕捉到图像中更大范围的上下文信息,而不增加计算成本;(2)空洞空间金字塔池化模块(ASPP)允许网络在多个尺度上捕获图像特征;(3)编码器-解码器结构用于精细化分割结果。编码器负责提取丰富的特征信息,解码器则将这些特征映射回原始图像尺寸,以产生精确的分割边缘。

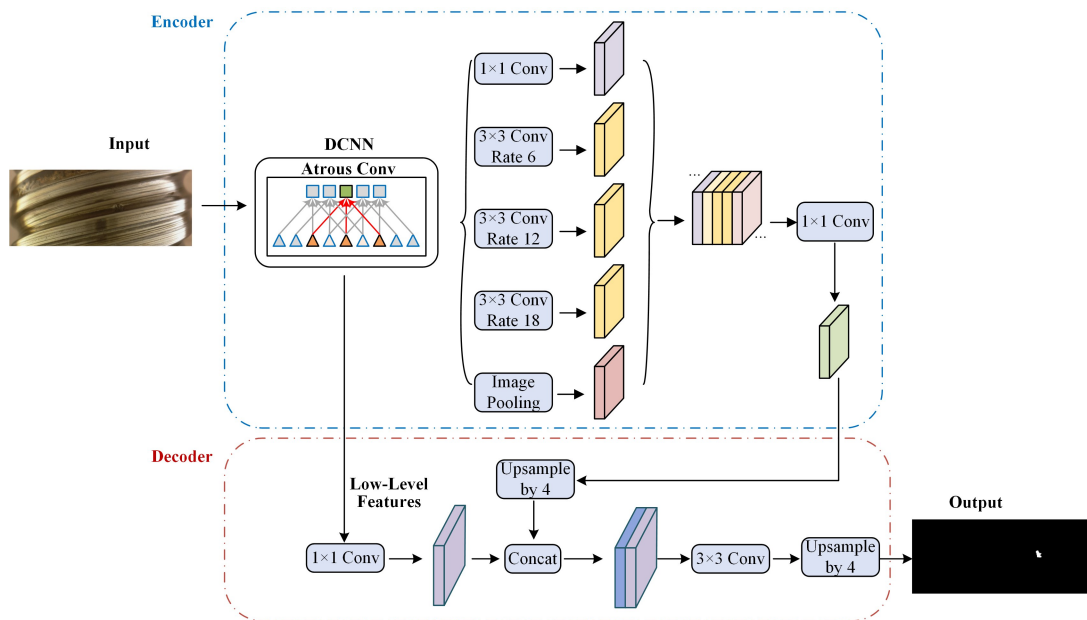


图 1 Deeplabv3+网络结构

Fig. 1 Deeplabv3+ network architecture

Deeplabv3+编码器阶段通常用 Xception 或 ResNet 作为其主干网络,用于特征提取,这一步通常会将特征图下采样 16 倍,加速模型的收敛速度。主干网络进行基础特征提取后,采用空洞空间金字塔池化(ASPP)模块通过结合不同空洞率[1,6,12,18]的空洞卷积层,捕获了更多尺度特征的特征信息,这对于处理图像中不同大小的对象尤为关键。解码器则需要将这些特征映射回原始图像大小,进行特征融合,然后得到最终预测的特征图进行像素级的预测。原始的 Deeplabv3+算法在自然场景中能够有效分割出目标区域,但其在编码器阶段直接将下采样 16 倍的特征图进行特征融合会丢失部分图像的细节信息,同时只使用空洞空间金字塔池化还不足以捕获足够的多尺度信息,尤其在缺陷在图像中占比较小情形下,会导致检测性能不佳。因此,针对 BSD 表面的点蚀缺陷检测,本文提出了一种改进的 Deeplabv3+算法。

3 改进的 DeepLabv3+ 网络结构

本文针对滚珠丝杠驱动(BSD)表面点蚀缺陷小且难以

检测的问题,提出了一种基于改进 Deeplabv3+模型的 BSD 表面缺陷检测方法。如图 2 所示,该方法的关键改进点包括:(1)使用 Res2Net-50 替换传统的 ResNet-50 主干网络,增强模型在内部提取丰富多尺度特征的能力;(2)在空洞空间金字塔池化(ASPP)后加入 Coordinate Attention 注意力机制,进一步优化特征图的空间和通道信息;(3)在编码器和解码器之间引入特征金字塔网络(FPN)进行多尺度特征的有效融合,克服了原始 Deeplabv3+模型在特征融合时可能导致的细节信息丢失问题。在主干网络部分,模型被分为 4 个阶段,即图中的 $1/2f$, $1/4f$, $1/8f$ 和 $1/16f$ 阶段,其中“f”表示特征图,前缀数字表示相对于原图的尺度大小。FPN 首先将 $1/16f$ 的特征图上采样 2 倍,与 $1/8f$ 的特征图进行融合,生成特征信息更丰富的 $1/8f$ 特征图。经 Coordinate Attention 处理的 $1/16f$ 特征图与融合后的 $1/8f$ 特征图再次进行 FPN 融合,以进一步增强特征表达。最终,在解码器阶段,将编码器阶段的 $1/4f$ 特征图送入解码器,并与前述融合特征拼接,得到含有丰富细节和语义信息的特征图。通过进一步的卷积和上采样操作,模型生成最终的预测特征图。

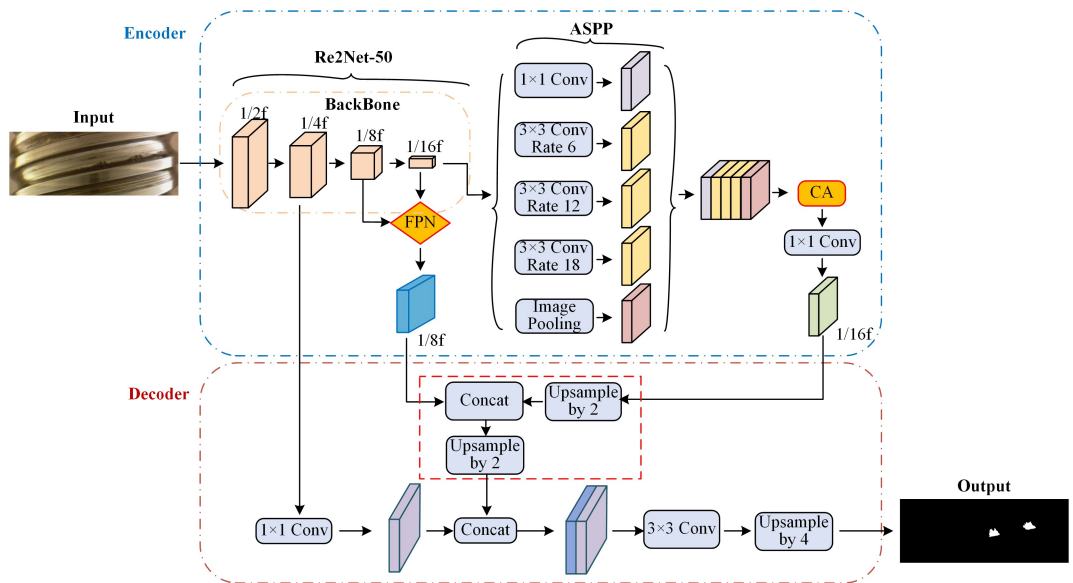


图2 改进的 Deeplabv3+ 网络结构

Fig. 2 Improved Deeplabv3+ network architecture

3.1 引入 Re2Net 骨干网络

改进的 DeepLabv3+ 网络以 Res2Net-50 为主干网络,如图 3 所示,Res2Net^[9]将每个残差块中的通道分成若干组,并在每组内进行独立卷积。假设一个特征图 X 的通道数为 C ,这个特征图将被分为 G 组,每组包含 C/G 个通道。对于每一组 $i(1 \leq i \leq G)$,其计算式可以表示为:

$$Y_i = F_i(Y_{i-1} \oplus X_i) \quad (1)$$

其中, F_i 表示第 i 组的卷积操作, Y_i 是 i 组的输出特征图, X_i 是输入特征图 X 的第 i 组通道, \oplus 表示特征图进行拼接的操作。最终所有分组的输出 Y_i 会被拼接起来,形成整个残差块的输出 Y ,可以表示为:

$$Y = Y_1 \oplus Y_2 \oplus \dots \oplus Y_G \quad (2)$$

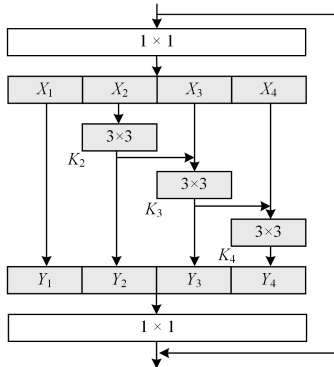


图3 Re2Net 模块

Fig. 3 Re2Net module

这种结构使得 Res2Net-50 能够在每个块内捕获丰富的多尺度特征,而整体网络深度和通道数与原始的 ResNet-50 保持一致。例如,各阶段通道数分别为 64, 128, 256 和 512,与原始的 ResNet-50 相同。这样的设计不仅增强了模型的特征表示能力,而且保持了相对较低的计算复杂度和参数量。

3.2 Coordinate attention 注意力机制

传统的 SENet^[15] 注意力网络主要专注于探索各通道之间的相互关系,以评估每个通道的重要性。然而,这种方法通常会忽略目标在图像中的空间位置信息,这是理解图像内容的关键要素之一。作为一种改进,CBAM^[16] 注意力机制在

SENet 的基础上增添了空间注意力模块。尽管如此,CBAM 中的通道注意力和空间注意力模块仍然是独立运作的,这在一定程度上限制了其性能的提升。基于此,本文引入了一种创新的注意力机制 Coordinate Attention,有效地整合了空间和通道信息。如图 4 所示,Coordinate Attention 的步骤包括:(1)首先对输入特征图进行分解,并沿着垂直和水平方向聚合特征以编码位置信息;(2)使用两个独立的卷积层处理这些编码的位置信息,生成分别对应垂直和水平方向的注意力图;(3)通过元素级加法操作融合这两个注意力图;(4)将融合后的注意力图应用于原始特征图。Coordinate Attention 通过这种方式,不仅考虑了通道间的关系,还充分利用了空间位置信息,为更准确和详尽的图像解析提供了新的视角。其核心原理表现为:

$$Y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (3)$$

其中, $g_c^h(i)$ 和 $g_c^w(j)$ 分别表示沿水平坐标和垂直坐标对每个通道进行编码, $x_c(i, j)$ 表示原始的特征图。

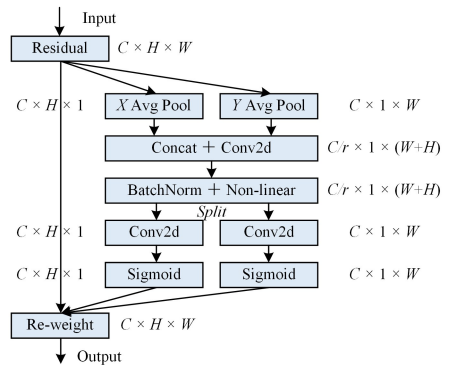


图4 Coordinate Attention 模块

Fig. 4 Coordinate Attention module

3.3 基于 FPN 的多尺度特征融合机制

FPN 通过结合不同层级的特征,解决了在单一尺度上进行特征提取时可能遇到的限制,使得模型能够同时捕捉到图像中的细节和语义信息,从而在处理复杂的视觉任务时表现出更高的性能。本文中主要使用了两层的金字塔特征融合方式,即将小尺度的特征图上采样 2 倍和较大的特征进行拼接

融合的操作。原始的 DeepLabV3+模型中,直接将下采样 16 倍的特征送入解码器可能不足以有效检测小目标缺陷。因此,本文提出在主干网络的 1/8f 和 1/16f 处加入 FPN 网络,以进一步提高特征的利用率和融合效果。同时,为了保持足够的细节信息,1/4f 阶段的特征图也会被保留并送到解码器。在此基础上,经 Coordinate Attention 处理的 1/16f 特征图与融合后的 1/8f 特征图再次进行 FPN 融合,以进一步增强特征表达。

通过这些综合性的改进,本文构建了一个经过精心设计的、改进版的 DeepLabV3+模型。该模型通过其各个模块的相互增强,极大地提高了对复杂场景中多尺度特征的捕获能力,尤其在精确分割微小缺陷和复杂形状方面表现出了卓越的性能。

4 实验结果与讨论

4.1 数据集与试验平台比较

本文的实验验证是在 Schlagenhauf 等提出的 BSD^[7]数据集上进行的。BSD 原始数据集包含 1 104 张图片,作者从中精心挑选了 394 张图片进行详细标注。这些图片中,324 张用于构建训练集,而剩余的 70 张则用于测试集。本文所有的实验均在 Pytorch 框架下完成,确保了实验的一致性和可重复性,平台参数如表 1 所列。同时为了保持数据的准确性,本文方法与所有的对比方法均在该数据集的训练集上进行训练,在测试集上进行测试。

表 1 实验平台参数

配置名称	参数
操作系统	Ubuntu
CPU	Intel(R) Xeon(R) Platinum 8368
GPU	RTX-A6000

4.2 评价指标

为了客观评价本文改进的 DeepLabV3+的性能,将平均交并比(MIoU)、准确率(Accuracy)、F1 得分(F1-Score)作为评价指标。其中 MIoU 是一种常用于评估图像分割任务的指标,它计算预测分割和真实分割之间的平均交集与并集的比例。Accuracy 计算所有预测中正确预测的比例。F1-Score 是精确率和召回率的调和平均,更能均衡地反映模型的性能。通过这些公式,可以从不同的角度评估模型的性能。

因为是二分类的分割任务,所以 MIoU 可以简化为单个类别的 IoU 的计算公式:

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

准确率是正确预测的数量与总预测数量的比值:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

F1-Score 是精确率和召回率的调和平均:

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

其中,TP 为模型正确预测为正例的样本数量,TN 为模型正确预测为负例的样本数量,FP 为模型错误预测为正例的样

本数量,FN 为模型错误预测为负例的样本数量。

4.3 消融实验

本文使用的 BSD 数据集的测试集涵盖了各种尺寸的点蚀缺陷,从几厘米大小的明显点蚀到几十毫米大小的难以肉眼分辨的缺陷。因此,该数据集能够全面反映模型在实际应用中的性能,而非仅仅反映过拟合所带来的性能提升。如表 2 所列,为了验证本文提出的各个模块在原始 Deeplabv3+网络基础上带来的性能提升,对添加的 Res2Net,FPN,Coordinate attention 3 个模块依次进行了消融对比实验。首先,从表中可以看出 Deeplabv3+模型在 BSD 测试集上的 Accuracy,F1-Score,MIoU 等核心评测指标上是最底的,分别为 88.57,90.75,84.37。随后,将骨干网络替换为 Re2Net 进行实验,Accuracy 指标提升到 91.26,F1-Score 指标提升到 91.79,MIoU 指标提升到 86.64,是 3 个模块中单独嵌入到原始 Deeplabv3+模型后提升最大的模块,这验证了在卷积内部进行多尺度特征交互和融合的有效性。3 个模块之间进行两两组合,验证不同模块组合的效果,实验结果表明提出的 3 个模块两两之间是相互增益的,不同模块的组合的指标相对于单独的模块都是有所提升的。Re2Net+CA 两个模块的嵌入效果相比其他方式的组合效果更佳,Accuracy,F1-Score,MIoU 分别达到了 93.17,93.08,87.79,与本文算法的指标最为接近。消融实验结果表明了本文提出的改进 Deeplabv3+模型各个模块的有效性,Re2Net 模块是增益最大的模块,Coordinate attention 其次,FPN 增益相对较小。

表 2 消融实验结果

方法	Accuracy	F1-Score	MIoU
Deeplabv3+	88.57	90.75	84.37
Deeplabv3+_Re2Net	91.26	91.79	86.64
Deeplabv3+_FPN	88.89	89.18	84.49
Deeplabv3+_CA	90.62	90.43	85.78
Deeplabv3+_Re2Net_FPN	91.79	92.44	86.82
Deeplabv3+_Re2Net_CA	93.17	93.08	87.79
Deeplabv3+_FPN_CA	90.48	91.49	86.02
本文方法	93.46	93.23	88.07

4.4 与主流分割算法比较

为了进一步说明本文改进的 Deeplabv3+算法在 BSD 点蚀缺陷识别上的优越性,将本文改进的算法和主流的 U-Net,Segformer,PSPnet,FastSCNN,Deeplabv3+,Segformer,SwinTransformer 等分割算法进行了对比(见表 3),加粗数据为评价指标最大值。

表 3 不同分割算法性能比较

Table 3 Comparison of performance of different segmentation algorithms

方法	Accuracy	F1-Score	MIoU
U-Net	87.42	89.54	82.69
Segformer	91.87	92.07	86.31
PSPnet	89.94	91.59	85.59
FastSCNN	85.51	85.21	77.15
Deeplabv3+	88.57	90.75	84.37
SwinTransformer	90.70	92.12	86.39
本文方法	93.46	93.23	88.07

由表 3 可知,FastSCNN 算法的指标在所有对比算法中是最底的,Deeplabv3+在所有对比算法中位居中间。本文改进的 Deeplabv3+算法的 MIoU 值为 88.07,F1-Score 为

93.32, Accuracy 为 93.46,是所有对比方法中指标最高的,说明了本文改进算法的有效性。同时本文也和基于 Transformer 架构的 Segformer 和 SwinTransformer 的分割网络进行了比较,算法在测试集的性能上同样处于领先地位,进一步说明了本文改进的 Deeplabv3+算法的有效性。

为了更直观地展示所提出的 Deeplabv3+算法的优越性,针对 MIoU 值,绘制了不同算法在不同迭代次数上的数值。由图 5 所示,本文改进算法为图中的红色曲线,可以看到红色曲线位于所有算法的上方,说明其 mIoU 值均高于对比算法,说明了本文改进的算法精度更高,性能更好。同样地,针对所比较的不同方法的 Accuracy 和 F1-Socre 指标也进行展示,如图 6 和图 7 所示,本文提出的算法的曲线位于对比方法的上方,进一步验证了本文方法的有效性。

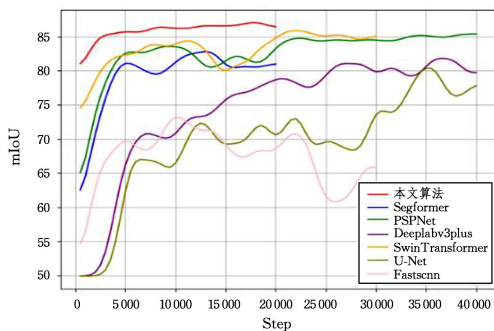


图 5 不同方法的 MIoU 性能对比图

Fig. 5 MIoU performance comparison of different methods

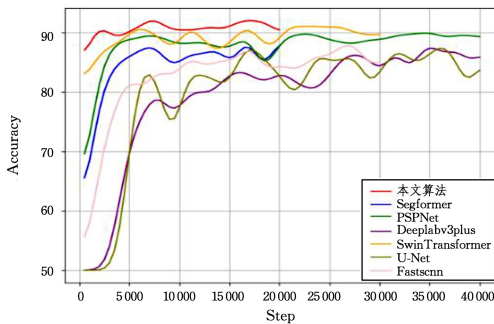


图 6 不同方法的 Accuracy 性能对比图

Fig. 6 Accuracy performance comparison of different methods

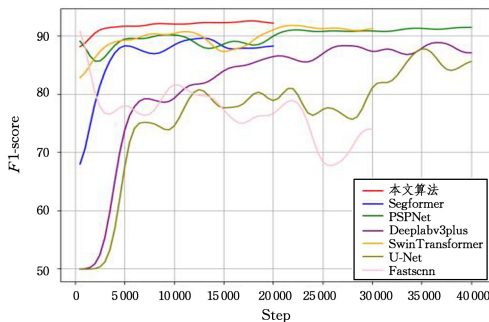


图 7 不同方法的 F1-Socre 性能对比图

Fig. 7 F1-Socre performance comparison of different methods

为了进一步展示改进算法的性能,给出了改进算法在 BSD 测试集上的可视化结果。测试集中的缺陷主要包括较明显的缺陷和细小的缺陷。因此,根据缺陷的特性,分别展示了较大缺陷和细小缺陷两组图片。其中图 8 展示了比较明显的缺陷,左侧表示原始的 BSD 图片,右侧展示了本文改进的

算法所检测到的结果。为了便于观察与分析,算法最终预测的掩码图像被映射到原始图像上进行展示。由图 8 可以看到,对于这种比较明显的点蚀缺陷,本文改进算法能够有效检测出来。图 9 展示了细小的缺陷图片示例,其缺陷位置,人眼都很难发现,因为这种细小的点蚀缺陷很不明显。图 9 右侧展示了本文提出的算法的效果,可以看到,其能够很好地将细小的缺陷检测出来。

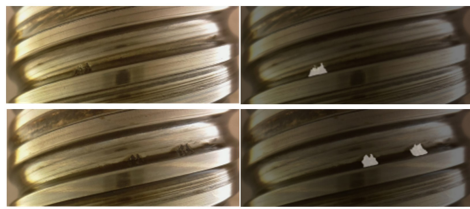


图 8 较大缺陷识别展示图

Fig. 8 Display of larger defect identification

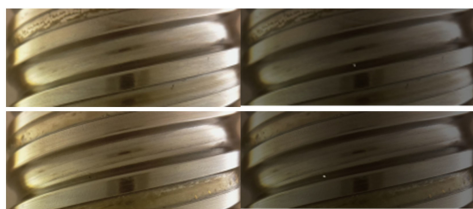


图 9 细小缺陷识别展示图

Fig. 9 Display of small defect identification

结束语 针对 Deeplabv3+模型在滚珠丝杠驱动点蚀小目标缺陷上分割的不足,本文采用 Re2Net-50 结构替换了原始的主干网络,同时在特征提取阶段针对语义信息丰富的特征图采用特征金字塔网络(FPN)进行特征融合,保留更丰富的语义信息。本文沿用了原始方法中的 ASPP 模块,并且针对小目标缺陷的特点对其进行了改进,在 ASPP 模块后面嵌入了 Coordinate Attention 注意力机制来弥补 ASPP 的不足,能够有效地融合空间与通道信息,提升提取的特征质量。实验结果表明,本文改进的 Deeplabv3+算法 MIoU 值为 88.07%,F1-Socre 为 93.32%,Accuracy 为 93.46%,能够满足实际生产中 BSD 缺陷检测的要求,尤其是细小缺陷的检测,能够有效地保证机床的安全运行。本文改进的 Deeplabv3+模型引入了 Res2Net,FPN 和 Coordinate Attention 等模块,这些模块的加入增加了模型的参数量和复杂性。在实际场景检测过程中,这可能导致检测速度变慢,因此在实际部署中存在一定的局限性。为了解决这一问题,未来的工作将着重研究轻量化网络模型,以提升模型的性能,以满足边缘端等实际场景的需求,并进一步提高模型的实用价值。

参考文献

- [1] LEISCHER F J, ROOS B A, CHOPP S M, et al. Lifecycle-oriented component selection for machine tools based on multibody simulation and component life prediction[J]. Journal of Manufacturing Science and Technology, 2009, 1(3): 179-184.
- [2] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.
- [3] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolu-

- tional networks for biomedical image segmentation[C]// Proceedings of the Medical Image Computing and Computer-Assisted Intervention. Munich, Germany, 2015; 234-241.
- [4] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017; 2881-2890.
- [5] PENG L, ZHANG H. U-net for Pavement Crack Detection[J]. Computer Science, 2021, 48(11A): 616-619.
- [6] LUO D L, CAI Y X, YANG Z H, et al. Survey on industrial defect detection with deep learning[J]. Science in China (Information Sciences), 2022, 52(6): 1002-1038.
- [7] SCHLAGENHAUF T, LANDWEHR M. Industrial machine tool component surface defect dataset[J]. Data in Brief, 2021, 39: 107643.
- [8] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]// Proceedings of the European Conference on Computer Vision (ECCV). 2018; 801-818.
- [9] GAO S H, CHENG M M, ZHAO K, et al. Res2Net: A New Multi-Scale Backbone Architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2): 652-662.
- [10] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017; 2117-2125.
- [11] HOU Q, ZHOU D, FENG J. Coordinate Attention for Efficient Mobile Network Design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021; 13708-13717.
- [12] PENG C, ZHANG K, MA Y, et al. Cross fusion net: A fast semantic segmentation network for small-scale semantic information capturing in aerial scenes[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60: 1-13.
- [13] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021; 10012-10022.
- [14] XIE E, WANG W, YU Z, et al. SegFormer: Simple and efficient design for semantic segmentation with transformers[J]. Advances in Neural Information Processing Systems, 2021, 34: 12077-12090.
- [15] JIE H, LI S, GANG S, et al. Squeeze-and-excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 8(42): 2011-2023.
- [16] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]// Proceedings of the European Conference on Computer Vision. 2018; 3-19.



LANG Lang, born in 1983, postgraduate, associate professor. His main research interests include machine learning, image processing, and anomaly detection.