

多线激光光条图像缺陷分割模型研究

郭晓轩¹, 冯其波^{1,3}, 冀振燕², 郑发家¹, 杨燕燕²

(1. 北京交通大学物理科学与工程学院, 北京 100044; 2. 北京交通大学软件学院, 北京 100044;
3. 东莞市诺丽科技股份有限公司, 广东东莞 523050)

摘要: 受环境干扰以及反射光影响, 室外采集的多线激光光条图像含有光斑和断裂缺陷. 为了准确地分割图像缺陷, 本文提出了一个轻量的UT(U-shape Target, U代表U型编解码网络结构, T代表靶形视野)分割模型, 模型由 3×3 卷积和靶形卷积堆叠而成. 靶形卷积是针对激光光条图像特点提出的多视野卷积模块, 模块中四个卷积分支构成靶形卷积视野, 能够提取激光光条图像几何结构特征、局部细节特征以及环绕纹理特征. 实验表明, UT模型在多线激光光条图像上的缺陷分割精度高于主流分割模型, 而且实现了分割精度和参数数量的平衡.

关键词: 缺陷分割; 激光图像; 深度学习; 轻量级分割模型; 多视野卷积

基金项目: 国家自然科学基金重点项目(No.51935002); 国家自然科学基金面上项目(No.52175493)

中图分类号: TP183

文献标识码: A

文章编号: 0372-2112(2023)01-0172-08

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20220644

Research on Segmentation Model of Multi-Line Laser Strip Image's Defects

GUO Xiao-xuan¹, FENG Qi-bo^{1,3}, JI Zhen-yan², ZHENG Fa-jia¹, YANG Yan-yan²

(1. School of Physical Science and Engineering, Beijing Jiaotong University, Beijing 100044, China;

2. School of Software, Beijing Jiaotong University, Beijing 100044, China;

3. Dongguan Nannar Technology Co., Ltd., Dongguan, Guangdong 523050, China)

Abstract: Influenced by environmental interference and reflected light, multi-line laser strip images collected outdoors contain the defects of flares and fractures. In order to segment the defects accurately, this paper proposes light-weight UT (U-shape Target, U represents a U-shaped encoder-decoder network architecture, and T represents a target-shaped receptive field) segmentation model, which stacks 3×3 convolutions and target convolutions. Considering the characteristics of laser strip images, we propose the target convolution, a multiple-receptive-field convolution module. Four convolution branches in this module form a target-shaped convolution receptive field, which can extract the geometric structure features, the local detail features and the surrounding texture features from the laser strip images. Experiments show that the UT model has higher defect segmentation accuracy than mainstream segmentation models, and can achieve the balance between the segmentation accuracy and the number of parameters.

Key words: defect segmentation; laser images; deep learning; light-weight segmentation model; multiple-receptive-field convolution

Foundation Item(s): National Natural Science Foundation of China (No.51935002, No.52175493)

1 引言

结构光视觉技术具有高精度、非接触的优点, 因此被广泛应用于工业检测领域, 包括轨道交通轮对故障检测领域. 轮对监测设备的面阵CCD相机和其对应的多线激光传感器阵列安装在轮对的左右两侧, 轮对经过定位传感器时, 左右两侧的CCD相机同时拍摄含有

多条激光线条的图片. 由于面阵CCD相机和其对应的多线激光传感器安装在室外环境, 成像过程受到环境干扰和反射光的影响, 导致图像含有光斑和断裂缺陷, 如图1所示. 图1中的光斑缺陷使多根光条粘连在一起, 光条断裂使形貌数据缺失, 这些成像缺陷严重影响了轮对的几何参数测量和形貌分析. 只有高精度地分

割这些缺陷位置才能对这些缺陷进行自动定位修复,进而准确地提取光条中心线进行高精度的轮对三维重建,最终解耦轮对几何参数进行轮对磨损检测和生命周期预测.考虑到轮对检测的精度和嵌入式设备部署要求,需要研究针对激光光条图像缺陷的高精度且轻量的分割模型.

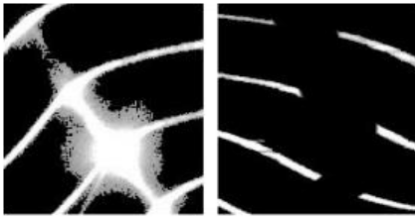


图1 光斑和断裂缺陷示意图

图像分割目前在智慧交通^[1]、医学图像分析^[2]、视频理解^[3]等领域得到广泛应用,其主要目的是根据灰度、颜色、纹理等特征将图像分成互不相交且具有各自特征的区域.近年来,基于深度学习的模型因具有自适应、鲁棒、精度高的优势获得广泛应用,其通过大量数据训练可自适应地定位目标区域的像素位置.部分模型通过构建自注意力机制^[4]、设计多分辨率分支^[5]、使用多视野卷积^[6,7]等获得了高精度的结果.文献[4]设计的自注意力机制能够根据物体区域特征表示与像素特征表示之间的相似度将这些物体区域特征表示传播给像素,进而得到增强的像素表示.文献[5]设计了多个不同分辨率的特征提取分支,通过多次分支间信息融合确保同时提取高级语义信息和空间信息.文献[6]模型首次证实网络宽度对精度有重要影响,通过设计多个卷积分支以提取丰富的多尺度特征.文献[7]设计了空洞空间卷积池化模块,模块由四个不同空洞率的卷积分支构成,可提取不同范围特征.

上述模型虽然获得了精确的分割结果,但是模型体积较庞大,难以部署在工业检测嵌入式设备上.工业检测嵌入式设备部署要求模型不仅精度高,还需轻量化.目前有很多轻量化模型设计方法,如堆叠轻量级卷积模块^[8,9]、解码器轻量设计^[10,11]、双分支架构^[12]、多尺度图像级联网络^[13]、简易U型模型搭建^[14,15]等.文献[8]在编码阶段设计了具有通道拆分和通道混洗操作的SS-nbt模块,减少参数数量的同时确保特征提取的有效性.文献[9]提出由两个不同分辨率分支构成的深度双分辨率网络,分支均采用轻量级残差Bottleneck模块减少参数量.文献[10]提出灵活的轻量级解码器(FLD),将语义特征图从高级别到低级别依次减少维度,降低了解码阶段的计算开销.文献[11]提出一种简易、轻量的解码器,采用 1×1 卷积提取特征,通过特征图相加融合不同阶段的解码特征,减少参数量且获得良好的解

码性能.文献[12]设计了由语义分支和空间分支构成的网络以提取高级语义信息和细节信息.语义分支采用窄通道、深层网络获取大感受野并捕获语义特征;空间分支采用宽通道、浅层网络提取低层特征图的细节信息.文献[13]设计了包含多重分辨率分支的图像级联网络,将低分辨率图输入到较深的卷积神经网络处理,将中高分辨率图输入到轻量的卷积神经网络处理,然后由级联特征融合单元和级联标签引导语义特征和细节特征的高效融合.文献[14]构建了U型结构的编解码网络,并设计跳跃连接机制向解码器传递编码信息.文献[15]设计了全尺寸跳跃连接的U型网络结构提高解码特征的多样性,并降低解码阶段的特征图维度以缩减参数量.

上述分割模型主要针对彩色图像设计,对于公共数据集具有良好的分割性能,但对于具有特殊几何结构的灰度多线激光光条图像的缺陷分割任务效果欠佳.为了满足工业缺陷检测领域高精度、低参数的需求,针对多线激光光条图像的特点,本文构建了一个轻量级的UT(U-shape Target)分割模型用于高精度地分割光条图像缺陷,UT模型的“U”代表模型的主体结构是U形,“T”代表模型的靶形卷积视野.

2 UT模型

2.1 UT模型结构

UT模型是一个U型结构的编解码网络,其结构如图2所示,模型由五层编码器和四层解码器构成,模型的核心是针对激光光条图像特点所设计提出的由四个卷积分支构成的靶形卷积,如图2中的虚线框所示,靶形卷积包含了横向条纹卷积分支、纵向条纹卷积分支、常规 3×3 卷积分支以及 3×3 空洞卷积分支.横向和纵向卷积分支通过大范围的矩形感受野提取条纹的几何结构特征,如完整光条几何信息、断裂几何信息、黏连几何信息等.常规 3×3 卷积分支感受野最小,主要提取局部区域的细节特征,如条纹的边缘像素点信息等. 3×3 空洞卷积分支有着环形感受野,主要提取环绕局部区域的纹理特征,为局部区域细节特征补充上下文信息.由于靶形卷积可以同时提取光条图像的几何结构特征、局部细节特征以及环绕纹理特征,因此可更准确地分割光条图像缺陷.

UT模型的五层编码器的特征图维度分别为16, 32, 64, 128和256;四层解码器的特征图维度均为80.UT模型中,各层编码器先采用 3×3 卷积提高特征图维度,再通过所提出的靶形卷积提取丰富的多视野特征.每层解码器均采用文献[15]的全尺寸跳跃连接获取细节信息和语义信息,然后使用靶形卷积进一步提取丰富的特征.在全尺寸跳跃连接中,每层编码特征通过

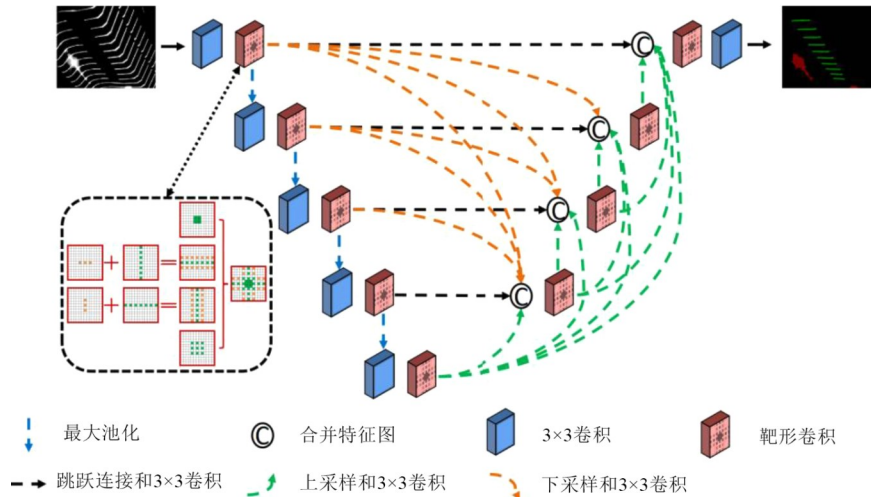


图2 UT模型结构

3×3卷积操作传递到相同分辨率的解码层,通过下采样和3×3卷积操作传递到更低分辨率的解码层;每层解码特征通过上采样和3×3卷积操作传递到更高分辨率的解码层.因此,每层解码器通过全尺寸跳跃连接组合了相同分辨率编码层特征、更高分辨率的编码层特征以及更低分辨率的解码层特征.来自高分辨率层的细节信息和低分辨率层的高级语义信息在各解码层组合,提高了解码特征的多样性.得益于全尺寸跳跃连接和靶形卷积的使用,UT模型不仅可以高精度地分割激光光条图像的缺陷,而且实现了精度和参数量的平衡.

2.2 靶形卷积模块

靶形卷积模块是UT模型的核心.传统的方形卷积视野不适合光条图像特征的提取,原因在于:为了提取丰富的光条几何结构特征需要大的卷积视野,而大视野的方形卷积会造成提取特征中冗余背景特征占比过大,影响分割效果.因此,本文针对多线激光光条图像的特点设计了靶形卷积视野,靶形卷积视野可更有效地提取更丰富的光条几何结构特征.得益于空洞卷积和一维卷积的使用,靶形卷积模块以较小的参数量获得了大的卷积视野.

2.2.1 靶形卷积感受野

多线激光光条图像与日常场景彩色图像不同,其为灰度图且包含大量狭长光条.图3展示了多线激光

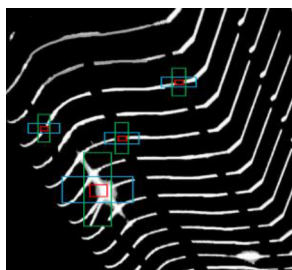


图3 多线激光光条图像与缺陷的感受野

光条图像,图中红绿蓝矩形构成的彩色图形代表了靶形卷积视野.从图3可以看出,由于光条结构特殊,如果使用传统的方形视野提取光条缺陷特征会受限,导致分割效果不佳.为了高效提取光条图像缺陷特征,首先设计由横向条纹视野(蓝色矩形框)和纵向条纹视野(绿色矩形框)构成的“十字形”视野.横向条纹视野关注单根光条内部结构,判断该区域是否存在断裂;纵向条纹视野关注相邻光条间的关联信息,判断该区域是否存在光条黏连(即光斑).因此,由横向和纵向条纹视野组合的“十字形”视野可以有效地提取光条的几何结构信息.

“十字形”卷积视野的原理如图4虚线方框所示,卷积由横向条纹卷积分支和纵向条纹卷积分支组成.横向条纹卷积分支首先使用3×1的空洞卷积提取较短的纵向特征,然后使用1×7的空洞卷积提取较长的横向特征;纵向条纹卷积分支首先使用1×3的空洞卷积提取较短的横向特征,然后使用7×1的空洞卷积提取较长的纵向特征.通过对两个一维空洞卷积的级联,横向条纹卷积分支和纵向条纹卷积分支分别获得一个更大范围的横向条带感受野和纵向条带感受野,进而组合成一个更大尺寸的“十字形”视野,从而有效地分割多线激光光条图像中较大尺寸的缺陷.

对于图像中包含的小尺寸光斑和小范围断裂,向“十字形”卷积添加两个3×3的方形卷积分支构建靶形卷积,其中传统的3×3卷积主要提取局部细节特征,3×3空洞卷积提取局部环绕特征.图4所示的四分支靶形卷积不仅可提取延展的几何结构特征,还可提取局部细节特征以及环绕特征,从而高效地分割激光光条图像中形状各异、大小不一的缺陷.

靶形卷积的计算方式如下:

$$M = H[\delta(B(\omega(x))), \delta(B(\mu(x))), \delta(B(\beta_v(\delta(B(\varphi_h(x)))))), \delta(B(\beta_h(\delta(B(\varphi_v(x))))))] \quad (1)$$

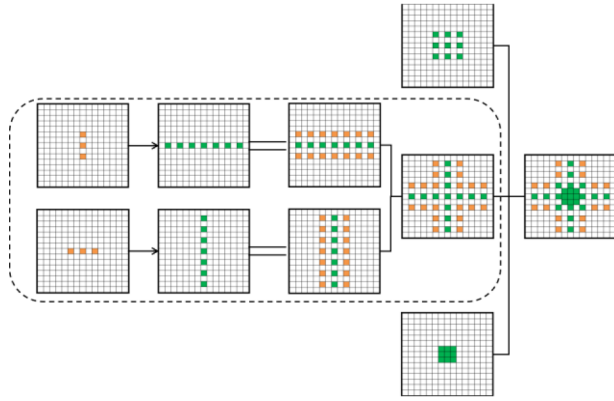


图4 靶形卷积的感受野

其中, x 表示输入特征图, M 表示经过靶形卷积操作后的输出特征图, ω 和 μ 分别表示传统的 3×3 卷积和 3×3 空洞卷积, B 表示批归一化操作, δ 表示 relu 激活函数, φ_h 和 φ_v 分别表示 1×3 空洞卷积和 3×1 空洞卷积, β_h 和 β_v 分别表示 1×7 空洞卷积和 7×1 空洞卷积, $H[\cdot]$ 表示特征图合并操作。

以 C_{in} 和 C_{out} 分别表示输入特征图(即 x) 和输出特征图(即 M) 的通道数; $\omega, \mu, \varphi_h, \varphi_v$ 的维度设置为 $C_{in} \times \frac{C_{out}}{4}$; β_h 和 β_v 的维度设置为 $\frac{C_{out}}{4} \times \frac{C_{out}}{4}$ 。

2.2.2 靶形卷积参数量

靶形卷积级联了一维卷积, 是一个轻量级的模块。靶形卷积和传统的 3×3 卷积的参数数量对比如下:

$$P_r = 3 \times 3 \times C_{in} \times C_{out} \quad (2)$$

$$P_t^r = 3 \times 3 \times C_{in} \times \frac{C_{out}}{4} \quad (3)$$

$$P_t^d = 3 \times 3 \times C_{in} \times \frac{C_{out}}{4} \quad (4)$$

$$P_t^v = 1 \times 3 \times C_{in} \times \frac{C_{out}}{4} + 7 \times 1 \times \frac{C_{out}}{4} \times \frac{C_{out}}{4} \quad (5)$$

$$P_t^h = 3 \times 1 \times C_{in} \times \frac{C_{out}}{4} + 1 \times 7 \times \frac{C_{out}}{4} \times \frac{C_{out}}{4} \quad (6)$$

$$P_t = P_t^r + P_t^d + P_t^v + P_t^h \quad (7)$$

$$R = \frac{P_t}{P_r} = \frac{48 \times C_{in} + 7 \times C_{out}}{72 \times C_{in}} \quad (8)$$

其中, P_r 和 P_t 分别表示常规的 3×3 卷积模块和靶形卷积模块的参数量; $P_t^r, P_t^d, P_t^v, P_t^h$ 分别表示靶形卷积中常规的 3×3 卷积分支、 3×3 空洞卷积分支、纵向条纹卷积分支以及横向条纹卷积分支的参数量; R 表示靶形卷积模块参数量同传统的 3×3 卷积模块参数量的比值。在 UT 模型的靶形卷积中, 输入维度等于输出维度(即 $C_{in} = C_{out}$), 因此 R 的数值为 $\frac{55}{77}$ (小于 1), 这表明 UT 模型的靶形卷积模块比传统的 3×3 卷积模块节省参数量。

3 实验

为了验证所提出的靶形卷积模块和 UT 模型在激光光条图像上的分割效果, 本文设计 4 组实验: (1) 不同下采样次数对比实验, 验证分割精度最高时的下采样次数; (2) 多视野卷积对比实验, 对比了靶形卷积和前沿的多视野卷积的分割精度; (3) 不同尺寸十字形卷积视野对比实验, 验证分割精度最高时的空洞率设置值; (4) 分割模型对比实验, 对比了 UT 模型与前沿分割模型的分割精度。

实验环境硬件配置为: Intel i9-10900X 处理器、NVIDIA GeForce RTX 3090 显卡。软件配置为: CUDA11.0+CuDnn8.0.5 并行计算框架、Python3.7 编程语言、Pytorch1.7.1 深度学习框架。

3.1 实验数据集

激光光条图像中的缺陷形态各异、尺寸不一, 且在图像中的分布位置也不固定。传统的图像分割模型难以应对如此复杂的分割场景, 因此需要采用基于深度学习的图像分割模型自适应且鲁棒地分割这些缺陷。基于深度学习的分割模型需要高质量的数据集进行训练, 目前没有用于多线激光光条图像缺陷分割的公共数据集存在, 因此需要构建数据集以训练所提出的分割模型。

用来构建数据集的多线激光光条图像均由安装在铁路轨道旁的轮对监测设备实地采集, 数据真实。为了模拟复杂的断裂效果, 本文在实采的轮对多线激光光条图像中多个位置擦除了部分正常光条, 然后通过差分减法获得断裂缺陷的具体像素位置。对于光斑, 由于光斑形状不规则难以人工标记, 采用同样方法, 先擦除光斑, 然后将擦除前和擦除后的图片进行差分减法以获得光斑的精确像素位置。图 5 展示了数据集中缺陷标记样例, 图中绿色部分代表断裂, 红色部分代表光斑。

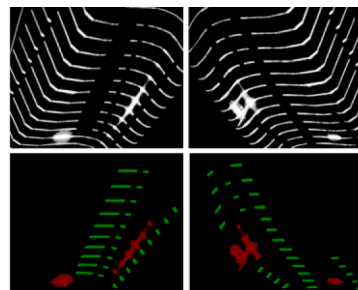


图5 缺陷标记示意图

所构建的多线激光光条图像缺陷分割数据集总共包含 2 766 对图像, 图像分辨率为 800×672 。

3.2 模型参数设置

实验将多线激光光条图像缺陷分割数据集中的

2 166 对用于模型训练,另外 600 对用于模型验证. 输入分辨率设置为 800×672 ,与数据集中图像分辨率保持一致. 模型的超参数设置见表 1.

表 1 模型超参数设置

参数	释义	取值
batch_size	每次迭代训练输入的图像数量	4
epoch	数据集全部样本训练完成的次数	60
optim	模型迭代训练的优化方法	Adam
lr	模型优化方法中的初始学习率	1×10^{-2}
loss	模型使用的损失函数	交叉熵

3.3 模型评价指标

针对轮对图像分割模型要同时满足精度和轻量化要求,实验采用 mIoU 和参数量作为评价指标,并将每个评价指标下最佳模型的结果值加粗显示. 参数量越小意味着模型占内存越少,适合嵌入式设备部署运行. mIoU (mean Intersection-over-Union) 是图像分割模型常用的精度衡量指标,数值越高意味着分割效果越好. 其中 IoU 表示某个分割类别的分割精度, mIoU 表示所有分割类别的精度平均值. mIoU 的计算方式如下:

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{n_{ii}}{\sum_{j=0}^k n_{ij} + \sum_{j=0}^k n_{ji} - n_{ii}} \quad (9)$$

其中, $k+1$ 表示需要分割的类别 (k 代表前景类别, 1 代表背景), n_{ij} 表示属于类别 i 的像素被错误分类为类别 j 的数量, n_{ii} 表示属于类别 i 的像素被正确预测的数量.

3.4 实验结果和分析

3.4.1 下采样次数对比实验

由于多线激光图像缺陷的尺寸差异较大,无法通过统计缺陷尺寸确定下采样次数. 为确定最佳下采样次数,将下采样次数分别设置为 2, 3, 4, 5, 通过实验对比 UT 模型分割精度. 实验结果如表 2 所示,下采样次数为 4 的 UT 模型 mIoU 最高,因此将 UT 模型下采样次数设置为 4.

表 2 不同下采样次数对比

模型	下采样次数	光斑 IoU/%	断裂 IoU/%	mIoU/%
UT	2	90.45	93.63	92.04
	3	90.42	94.32	92.37
	4	90.45	94.52	92.49
	5	90.25	94.56	92.41

3.4.2 多视野卷积对比实验

为了验证靶形卷积的分割效果,本文设计实验对比靶形卷积与其他多视野卷积. 所对比的多视野卷积如下:

(1) Inceptionv1: 文献[6]中具有维度缩减操作的 inception 模块,包含四个特征提取分支,可关注不同感

受野的信息.

(2) ASPP: 文献[7]中的基于空洞卷积的金字塔池化模块,含有四个不同空洞率的卷积分支,在没有增加参数的情况下扩大了感受野.

(3) RFB: 文献[16]中的感受野模块有三个特征提取分支,且每个分支级联了不同尺寸和空洞率的卷积模块,扩大了感受野并且避免了网格效应^[17].

(4) CG: 文献[18]中的特征提取模块,组合了常规的 3×3 卷积和 3×3 空洞卷积,可同时提取细节特征以及细节附近的环绕特征.

(5) AC: 文献[19]中的非对称卷积模块,具有三个卷积分支,分别为 3×3 卷积分支、 1×3 卷积分支和 3×1 卷积分支.

(6) 十字卷积: 所提出的十字形感受野卷积模块.

(7) 靶形卷积: 所提出的靶形感受野卷积模块.

实验中,将 unet3+ 模型中各编码器、解码器的最后一个 3×3 卷积模块分别用上述多视野卷积模块替换,以对比不同多视野卷积的分割效果,实验结果如表 3 所示.

表 3 靶形卷积和其他多视野卷积的对比

多视野卷积模块	基准模型	参数量/M	光斑 IoU/%	断裂 IoU/%	mIoU/%
无	Unet3+	1.69	90.04	94.30	92.17
Inceptionv1	Unet3+	1.03	90.33	94.11	92.22
ASPP	Unet3+	1.47	90.29	93.84	92.07
RFB	Unet3+	1.57	90.58	94.17	92.38
CG	Unet3+	1.69	89.97	94.35	92.16
AC	Unet3+	1.25	90.47	94.08	92.28
十字卷积	Unet3+	1.41	90.42	94.37	92.40
靶形卷积	Unet3+	1.45	90.45	94.52	92.49

从表 3 可以看出,靶形卷积的断裂 IoU 和 mIoU 最高,分别为 94.52% 和 92.49%,且参数量较低 (1.45 M). RFB 模块和 AC 模块的光斑 IoU 虽略微高于靶形卷积,分别超过 0.13% 和 0.02%,但断裂 IoU 分别低于靶形卷积 0.35% 和 0.44%. Inceptionv1 卷积模块虽然参数量最低,仅为 1.03 M,但在光斑和断裂的分割精度均不如靶形卷积. CG 模块、ASPP 模块、十字卷积模块的缺陷分割效果均不及靶形卷积. 综上所述,靶形卷积对多线激光光条图像缺陷的分割获得了最高的整体分割精度,且参数量较小,达到了分割精度和轻量化的综合最优,验证了靶形卷积视野设计的有效性.

3.4.3 不同尺寸十字形卷积视野对比实验

为了验证靶形卷积中不同尺寸的十字形卷积视野对多线激光光条图像缺陷分割精度的影响,将横向条纹和纵向条纹卷积的空洞率分别设置为 1, 2, 4 和 6,实验结果如表 4 所示.

表 4 中,UC 和 UT 分别表示具有十字卷积视野和靶形卷积视野的 U 型编解码网络. 实验结果显示,分割精度和十字形卷积视野的尺寸不成正比,其中空洞率为 2 时 UC 模型的分割精度达到最高,UT 模型的断裂分割精度和缺陷平均分割精度也达到最高. 因此,将模型横向条纹和纵向条纹卷积的空洞率设置为 2.

表 4 不同尺寸十字形卷积视野的分割精度对比

模型	空洞率	光斑 IoU/%	断裂 IoU/%	mIoU/%
UC	1	90.39	94.37	92.38
	2	90.42	94.37	92.40
	4	90.41	94.28	92.35
	6	90.01	94.00	92.01
UT	1	90.29	94.36	92.33
	2	90.45	94.52	92.49
	4	90.61	94.17	92.39
	6	90.25	94.43	92.34

3.4.4 分割模型对比实验

为了进一步验证 UT 模型的分割精度和轻量化,本文设计实验对比了 UT 模型和其他 16 个主流分割模型的参数量和分割精度. 这 16 个分割模型分别为 Unet^[14], Unet++^[20], AGUnet^[21], Unet3+^[15], Lednet^[8], Dfanet^[22], ICnet^[13], Enet^[23], CGnet^[18], EspnetV2^[24], BisenetV1^[25], BisenetV2^[12], OCRNet^[4], PP-LiteSeg^[10], DDRNet^[9]和 RegSeg^[11]. 实验结果如表 5 所示.

表 5 显示,本文所提出的 UT 模型和 Unet 系列的分割模型 Unet, Unet++, Unet3+, AGUnet 均获得了高于其他模型的缺陷分割精度,验证了所提出 UT 模型的 U 型结构的合理性. UT 模型的断裂分割精度、缺陷平均分割精度均超过其他 16 个主流分割模型,仅光斑分割精度落后于 AGUnet 0.01%. 模型 Enet, CGnet, Lednet 虽然参数量比 UT 模型少,但光斑和断裂分割精度均低于 UT

表 5 分割模型的对比

模型	参数量/M	光斑 IoU/%	断裂 IoU/%	mIoU/%
Unet	2.16	90.11	94.29	92.20
Unet++	2.55	90.15	94.26	92.21
AGUnet	2.19	90.46	94.08	92.27
Unet3+	1.69	90.04	94.30	92.17
Lednet	0.92	84.55	92.03	88.29
Dfanet	2.10	85.56	91.89	88.73
ICnet	26.24	76.06	82.41	79.24
Enet	0.35	89.06	93.05	91.06
CGnet	0.49	84.30	92.09	88.20
EspnetV2	2.16	86.70	91.16	88.93
BisenetV1	13.29	89.19	93.10	91.15
BisenetV2	3.54	86.13	93.37	89.75
OCRNet	70.35	86.18	93.54	89.86
PP-LiteSeg	11.27	84.70	92.05	88.38
DDRNet	5.70	84.80	92.28	88.54
RegSeg	3.33	86.28	93.56	89.92
UT	1.45	90.45	94.52	92.49

模型,缺陷平均分割精度分别低于 UT 模型 1.43%, 4.29% 和 4.2%. UT 模型对比 Unet 系列分割模型虽然精度提升有限,但是参数量均低于这些轻量级模型,以最低的参数量(仅 1.45 M)获得了最高的缺陷平均分割精度,UT 分割模型的参数量是 Unet3+ 模型的 85.8%、Unet 模型的 67.1%、AGUnet 模型的 66.2%、Unet++ 模型的 56.9%. UT 模型达到了高精度与轻量的综合最优,满足轮对检测对分割模型的要求.

为了更直观地观察各模型的分割效果,本文随机选取四张含有缺陷的多线激光图片以对比各模型的缺陷分割可视化效果,如图 6 所示. 第一列为来源于数据集的输入缺陷图,由于图片分辨率太大(800×672),其

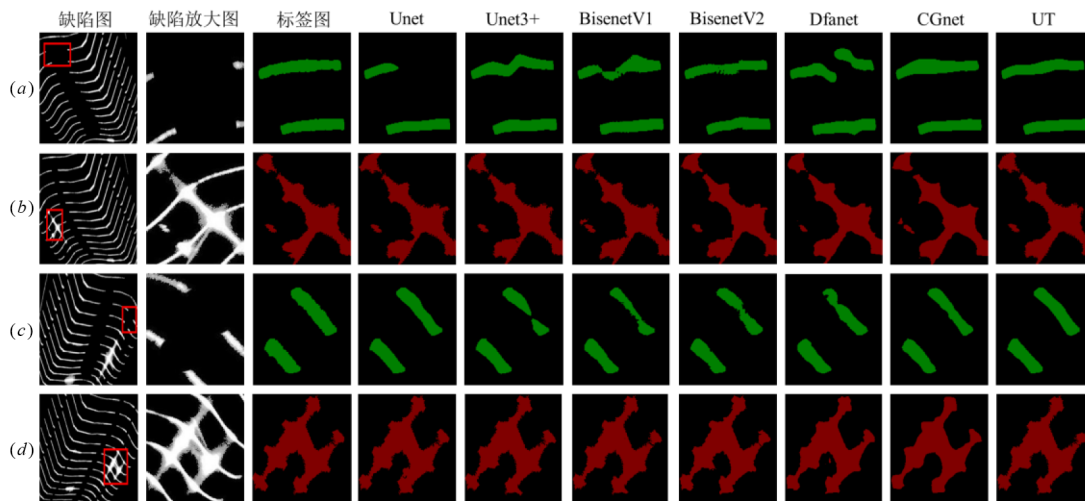


图 6 不同分割模型的可视化示意图

分割可视化图无法清晰展示缺陷边缘的分割效果,为此本文放大了各模型在缺陷区域的分割可视化结果.图中(a)、(b)、(c)、(d)行分别展示了各模型在直线断裂区域、小面积光斑区域、曲线断裂区域、大面积光斑区域的分割效果.

如图6所示,Unet不能完整地分割直线断裂区域,在小光斑缺陷区域则分割了过多的光条.Unet3+在直线断裂区域分割的轮廓不光滑,在曲线断裂区域分割轮廓不完整,在大光斑区域分割了过多的光条.BisenetV1在直线断裂区域分割的轮廓不光滑,在曲线断裂区域分割轮廓过细.BisenetV2在直线断裂和曲线断裂处分割的轮廓不光滑.Dfanet在直线断裂区域分割的轮廓有断裂,在光斑区域分割的轮廓边缘粗糙.CGnet在光斑区域分割的轮廓边缘也粗糙.对比之下,UT模型在这四种缺陷区域均能分割出完整的轮廓以及精细的轮廓边缘.

4 结论

为了分割多线激光光条图像缺陷以高精度地提取光条中心线,本文针对光条图像缺陷特点提出了由四个卷积分支构建的靶形视野,提取光条图像中丰富的几何结构信息以鲁棒地应对形态各异、大小不一的成像缺陷,并进一步构造了轻量的UT分割模型,模型在解码阶段能够融合全尺寸的多视野特征从而提高解码特征的多样性.为了训练和验证模型,本文构建了多线激光图像缺陷分割数据集.实验验证了靶形卷积和UT模型对多线激光光条图像缺陷的分割效果优于主流多视野卷积模块和分割模型,且UT模型满足嵌入式设备的轻量要求.由于激光光条图像缺陷标记工作耗时耗力,未来我们将进行半监督或自监督的激光光条图像缺陷分割模型的研究.

参考文献

- [1] 徐频捷,陈逸杰,李之南,等.基于事件驱动的车道线识别算法研究[J].电子学报,2021,49(7):1379-1385.
XU P J, CHEN Y J, LI Z N, et al. Research on event-driven lane recognition algorithms[J]. Acta Electronica Sinica, 2021, 49(7): 1379-1385. (in Chinese)
- [2] 赖小波,许茂盛,徐小媚.多分类CNN的胶质母细胞瘤多模态MR图像分割[J].电子学报,2019,47(8):1738-1747.
LAI X B, XU M S, XU X M. Glioblastoma multiforme multi-modal MR images segmentation using multi-class CNN[J]. Acta Electronica Sinica, 2019, 47(8): 1738-1747. (in Chinese)
- [3] 付利华,赵宇,姜涵煦,等.基于前景感知视觉注意的半

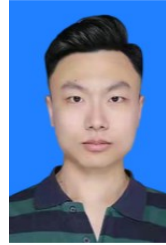
- 监督视频目标分割[J].电子学报,2022,50(1):195-206.
FU L H, ZHAO Y, JIANG H X, et al. Semi-supervised video object segmentation based on foreground perception visual attention[J]. Acta Electronica Sinica, 2022, 50(1): 195-206. (in Chinese)
- [4] YUAN Y H, CHEN X L, WANG J D. Object-contextual representations for semantic segmentation[C]//European Conference on Computer Vision. Glasgow: Springer, 2020: 173-190.
- [5] SUN K, XIAO B, LIU D, et al. Deep high-resolution representation learning for human pose estimation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 5686-5696.
- [6] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 1-9.
- [7] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [8] WANG Y, ZHOU Q, LIU J, et al. Lednet: A lightweight encoder-decoder network for real-time semantic segmentation[C]//2019 IEEE International Conference on Image Processing. Taipei: IEEE, 2019: 1860-1864.
- [9] HONG Y D, PAN H H, SUN W C, et al. Deep dual-resolution networks for real-time and accurate semantic segmentation of road scenes[EB/OL]. (2021-01-15) [2022-06]. <https://arxiv.org/abs/2101.06085>.
- [10] PENG J C, LIU Y, TANG S Y, et al. PP-LiteSeg: A superior real-time semantic segmentation model[EB/OL]. (2022-04-06) [2022-06]. <https://arxiv.org/abs/2204.02681>.
- [11] GAO R. Rethink dilated convolution for real-time semantic segmentation[EB/OL]. (2021-11-18) [2022-06]. <https://arxiv.org/abs/2111.09957>.
- [12] YU C Q, GAO C X, WANG J B, et al. BiSeNet V2: Bilateral network with guided aggregation for real-time semantic segmentation[J]. International Journal of Computer Vision, 2021, 129(11): 3051-3068.
- [13] ZHAO H S, QI X J, SHEN X Y, et al. ICNet for real-time semantic segmentation on high-resolution images[C]//European Conference on Computer Vision. Munich: Springer, 2018: 418-434.
- [14] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Comput-

ing and Computer-Assisted Intervention. Munich: Springer, 2015: 234-241.

- [15] HUANG H M, LIN L F, TONG R F, et al. UNet 3: A full-scale connected UNet for medical image segmentation [C]//ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing. Barcelona: IEEE, 2020: 1055-1059.
- [16] LIU S T, HUANG D, WANG Y H. Receptive field block net for accurate and fast object detection[C]//European Conference on Computer Vision. Munich: Springer, 2018: 404-419.
- [17] WANG P Q, CHEN P F, YUAN Y, et al. Understanding convolution for semantic segmentation[C]//2018 IEEE Winter Conference on Applications of Computer Vision. Lake Tahoe: IEEE, 2018: 1451-1460.
- [18] WU T Y, TANG S, ZHANG R, et al. CGNet: A light-weight context guided network for semantic segmentation [J]. IEEE Transactions on Image Processing, 2021, 30: 1169-1179.
- [19] DING X H, GUO Y C, DING G G, et al. ACNet: Strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 1911-1920.
- [20] ZHOU Z W, RAHMAN SIDDIQUEE M M, TAJBAKHSN, et al. UNet++: A nested U-Net architecture for medical image segmentation[C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Granada: Springer, 2018: 3-11.
- [21] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-net: Learning where to look for the pancreas[EB/OL]. (2018-04-11)[2022-06]. <https://arxiv.org/abs/1804.03999>.
- [22] LI H C, XIONG P F, FAN H Q, et al. DFANet: deep feature aggregation for real-time semantic segmentation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 9514-9523.
- [23] PASZKE A, CHAURASIA A, KIM S, et al. ENet: A deep neural network architecture for real-time semantic segmentation[EB/OL]. (2016-06-07) [2022-06]. <https://arxiv.org/abs/1606.02147>.
- [24] MEHTA S, RASTEGARI M, SHAPIRO L, et al. ESPNetv2: A light-weight, power efficient, and general purpose convolutional neural network[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 9182-9192.
- [25] YU C Q, WANG J B, PENG C, et al. BiSeNet: Bilateral

segmentation network for real-time semantic segmentation[C]//European Conference on Computer Vision. Munich: Springer, 2018: 334-349.

作者简介



郭晓轩 男, 1996年4月出生, 山东人. 北京交通大学物理科学与工程学院博士研究生. 主要研究方向为机器视觉和计算机视觉.
E-mail: 20118037@bjtu.edu.cn



冯其波(通讯作者) 男, 1962年5月出生, 广东人. 博士. 北京交通大学物理科学与工程学院教授, 博士生导师. 研究方向为光学测量、铁路安全测量技术、仪器仪表、机器视觉等.
E-mail: qbfeng@bjtu.edu.cn



冀振燕 女, 1972年4月出生, 河南人. 博士. 北京交通大学软件学院副教授, 博士生导师. 主要研究方向为计算机视觉、多源异构数据融合等.
E-mail: zhyji@bjtu.edu.cn



郑发家 男, 1991年2月出生, 安徽人. 2021年于北京交通大学获得博士学位, 现为北京交通大学讲师. 主要研究方向为数控机床几何误差测量、轮对踏面几何参数在线测量、机器视觉.
E-mail: zhfajia@bjtu.edu.cn



杨燕燕 女, 1986年3月出生, 河南人. 北京交通大学讲师. 主要研究方向为机器学习、不确定性人工智能.
E-mail: yangyy@bjtu.edu.cn