DOI:10.16136/j.joel.2023.05.0289

基于改进 DeepLabV3+的铣床碎屑图像分割方法

张 闯1,刘秀平1*,袁 皓1,冯国栋1,闫焕营2

(1. 西安工程大学 电子信息学院,陕西 西安 710048; 2. 深圳罗博泰尔机器人有限公司,广东 深圳 518109)

摘要:针对铣床碎屑形状不规则导致图像分割中碎屑轮廓不清晰、分割精度低的问题,本文提出 一种改进的 DeepLabV3+铣床碎屑分割算法。首先在 DeepLabV3+的 Xcepetion 模块中嵌入通道 与空间注意力机制(convolutional block attention module, CBAM)模型,优化通道的权重和位置信息,加强碎屑图像区域的特征学习;其次将 DeepLabV3+的空洞空间卷积池化金字塔(atrous spatial pyramid pooling, ASPP)模块改为密集连接(dense conclutional network, DenseNet)方式,增大碎 屑图像特征点的感受野,提升铣床碎屑图像特征的复用效率;最后在解码过程中采用多尺度自适 应特征融合方法,聚合多尺度特征作为解码器的输入特征,提高碎屑图像分割的精度与鲁棒性。 实验结果表明,本文算法优于其他分割算法,改进后算法相比 DeepLabV3+,像素准确率提高 0.026,平均交并比(mean intersection over union, *MIOU*)提高 0.020, *F*₁值提高了 0.013。 关键词:图像处理;铣床碎屑; DeepLabV3+; CBAM; 多尺度融合 **中图分类号:**TP391 **文献标识码:**A **文章编号:**1005-0086(2023)05-0489-09

Segmentation method of milling machine debris image based on improved DeepLabV3+

ZHANG Chuang¹, LIU Xiuping^{1*}, YUAN Hao¹, FENG Guodong¹, YAN Huanying² (1. School of Electronics and Information Xi'an Polytechnic University, Xi'an, Shaanxi 710048, China; 2. Municipal Robotel Robot Technology Co, LTD, Shenzhen, Guangdong 518109, China)

Abstract: Aiming at the problem that the irregular shape of the milling machine debris leads to the unclear outline of the debris and the low segmentation accuracy in the image segmentation, this paper proposes an improved DeepLabV3 + milling machine debris segmentation algorithm. First, the CBAM model is embedded in the Xcepetion module of DeepLabV3+ to optimize the weight and position information of the channel, and to strengthen the feature learning of the debris image area. Secondly, the AS-PP module of DeepLabV3+ is changed to the dense connection method to increase the receptive field of the feature points of the debris image. The feature reuse efficiency of the milling machine debris image is improved. Finally, the multi-scale adaptive feature fusion method is adopted in the decoding process, and the multi-scale features are aggregated as the input features of the decoder to improve the segmentation accuracy and robustness of the debris image. The experimental results show that the algorithm in this paper is better than other segmentation algorithms. Compared with DeepLabV3+, the improved algorithm improves the pixel accuracy by 0.026, mean intersection over union (*MIOU*) by 0.020, and the F_1 -measure by 0.013.

Key words: image processing; milling machine debris; DeepLabV3+; CBAM; multi-scale fusion

0 引 言

随着机器视觉在工业领域的发展,图像分割

作为一种新的目标检测手段被广泛应用^[1]。为了 实现工业智能化,在生产加工领域对铣床碎屑量 的实时监控显得尤为重要,因此如何准确对铣床 碎屑图像进行分割处理,成为研究该问题的首要 目标。图像分割就是根据图像中目标的某些特 征,如像素灰度值、颜色纹理以及形状轮廓等,将 图像分解成多个互不相交的区域,分解后同一区 域内图像具有相同特征[2]。近年来,随着计算机 硬件的发展,研究人员为赋予分割区域更精准的 语义信息[3],将之前处于理论阶段的深度学习引 入图像分割实践中,从而在图像分割方面取得技 术性突破^[4],如全卷积网络^[5](full convolutional network, FCN)、金字塔场景解析网络^[6] (pyramid scene parsing network, PSPNet)、DeepLab 系列^[7] 等基于深度学习的图像分割法。FCN 作为深度学 习在图像分割领域的首个算法,不仅对目标进行 像素级别的图像分割,并创建图像分割通用的网 络模型。但在 FCN 网络模型中,图像的上采样结 果较为模糊,并且对图像进行语义分割时,没有充 分考虑特征图中像素间的关系,导致分割结果缺 乏空间一致性^[8]。PSPNet 通过使用空间金字塔 池化的方法将图像局部与全局特征融合到一起, 使其在多个应用场景中的分割效果都优于 FCN。 但 PSPNet 在目标之间有遮挡的场景下,会丢失目 标边缘信息,使得目标分割的轮廓缺失。DeepLab 系列算法的提出大幅度优化了 PSPNet 在这一方 面的不足, DeepLab 系列算法在 FCN 基础上引入 空洞卷积,使得该算法在计算量不变情况下,感受 野增大,因此该算法在工业图像分割领域的适应 性较广。在金属分割领域, DENG 等^[9]提出了 Sea-thru 和 Mask R-CNN (mask region-based convolutional neural network)相结合的金属结核图像 分割模型,该方法使用 Sea-thru 算法对不同光照 的图像进行预处理提高图像细节,然后使用 Mask R-CNN 算法进行分割。但该算法在金属存在遮 挡情况下,表现不佳。YUAN等^[10]提出了基于 VGG(visual geometry group)的整体嵌套边缘检测 模型进行分割,用于处理相互粘连的目标,但当目 标较小时,由于存在重叠度高,相似度高等问题, 出现欠分割。

本文根据铣床碎屑的不规则轮廓、铣床碎屑 加工环境中受到光照强度不同等因素造成原始 DeepLabV3+网络分割精度差的问题,对原网络 进行如下改进。首先针对碎屑在复杂光照下特征 提取不佳情况,引入通道与空间注意力机制(convolutional block attention module, CBAM),以增强 铣床碎屑在复杂背景中的显著度和有用特征的提 取能力;其次通过密集连接(dense convolutional network, DenseNet)空洞空间卷积池化金字塔(atrous spatial pyramid pooling, ASPP)方式,增强 AS- PP内各个碎屑特征图的相关性,提高特征信息的 复用率;最后针对 DeepLabV3+解码过程中对编 码阶段特征信息利用不足的问题,引入自适应空 间融合(adaptively spatial feature fusion, ASFF),将 多级特征自适应融合后用于解码,提高图像分割 的精度。

1 DeepLabV3+网络模型

DeepLabV3+引入编码一解码(Enconder-Decoder)模块^[11]。该网络将 DeepLabV3 的网络作为 编码模块,在编码部分首先将图像输入网络进行一 系列卷积,其中主干网络为含有输入流(entry flow)、 中间流(middle flow)和输出流(exit flow)的 Xception^[12]网络,在进行碎屑特征提取的时会产生 1/2、 1/4、1/8和1/16大小的碎屑特征图像,不同大小的 碎屑特征图包含了不同尺度的上下文信息,最后将 得到的 1/16 大小的碎屑特征图传入 ASPP 模块。 ASPP 模块将输入的铣床碎屑特征图以并联的方式 分别进行 1×1 卷积、空洞率大小为 6、12、18 的 3×3 卷积和平均池化操作,然后将得到的铣床碎屑特征 图进行级联操作,最后再次经过1×1卷积传入解码 模块。在解码模块中,首先将经过特征提取主干网 络 Xception 输出的碎屑特征图进行 1×1 卷积,然后 与编码器输出碎屑特征图的4倍上采样进行级联, 再经过4倍上采样恢复为碎屑图像标准尺寸,最后 解码出预测碎屑图。由于该网络的编码模块在训练 数据集过程中会减少输入碎屑图像的空间维度,从 而导致部分碎屑的有用信息丢失,影响最终分割的 预测结果。

2 改进 DeepLabV3+网络模型

在 DeepLabV3+原网络上进行碎屑分割实验, 发现分割效果较差。碎屑分割图中存在碎屑边缘不 平滑现象,且鲁棒性较差,对于小目标容易出现漏分 割现象。由于在编码模块中参考了多种尺度的碎屑 特征信息,并在学习更深层的图像特征时加入池化 操作,则会导致在分割碎屑图像时部分碎屑的边界 轮廓结构信息丢失,比如碎屑边缘的一些细节和简 单的几何图形信息。为了改进由于主干网络逐层卷 积下采样导致的碎屑特征信息丢失问题,提升网络 对铣床碎屑的分割性能,本文对 DeepLabV3+网络 的改进如下所示。

首先,将 CBAM 模块嵌入到主干网络 Xception 的输入流、中间流和输出流,强化有效特征提取过 程,来提高网络的鲁棒性;其次,在编码阶段,本实验 采用 DenseNet 的 ASPP 模块,通过跨层连接的方 式,将上面所有层输出的碎屑特征图融合作为当前 层的输入特征图,提高特征传播效率;最后,引入 AS-FF 网络,通过融合多尺度特征图所含的丰富碎屑边 缘特征信息,来提高分割精度。改进后的网络如图 1 所示。



图 1 改进 DeepLabV3+网络模型 Fig. 1 Improving DeepLabV3+ network model

2.1 多注意力机制

铣床碎屑的特征较为复杂,其轮廓边界难以明确分辨,因此通过嵌入 CBAM 注意力机制来增加有用特征的权重。CBAM 注意力机制是一种轻量级的注意力模型,由通道注意力模块和空间注意力模块

两部分串联构成。前者主要保持通道维度不变,压 缩空间维度,主要关注输入图片中有意义的特征。 后者是空间维度不变,压缩通道维度,主要关注特征 的位置信息^[13]。CBAM 的通道和空间注意力网络 结构如图 2 所示。



图 2 空间与通道注意力机制 Fig. 2 Space and channel attention mechanism

通道注意力机制模块首先输入大小为 $H \times W \times C$ 的特征F,其中特征的长、宽和通道数量分别为H、W 和 C。首先特征F经过两个并行的最大池化(*MaxPool*)和平均池化(*AvgPool*)对空间维度进行压缩,得到两个 $1 \times 1 \times C$ 特征图,分别记为 F_{avg}^{c} 和 F_{max}^{c} ,其次将特征图经过共享权值的多层感知器

(muti-layer percetron, *MLP*),最后将 *MLP* 运算后 相加的特征经过 sigmoid 激活函数得到权重系数 M_c $\in \mathbf{R}^{1 \times 1 \times C}$,输入特征 $\mathbf{F} = \mathbf{M}_c$ 相乘结果为通道注意力 处理的新特征^[14]。具体计算如下式:

$$M_{c}(\mathbf{F}) = \sigma(MLP(AvgPool(\mathbf{F})) + MLP(MaxPool(\mathbf{F}))) =$$

 $\sigma(W_1(W_0(F_{avg}^{\epsilon})) + W_1(W_0(F_{max}^{\epsilon}))),$ (1) 式中, σ 指 sigmoid 函数, MLP(AvgPool(F))指全局 平均池化(global average pooling, GAP)的权重, MLP(MaxPool(F))指全局最大池化(global max pooling, GMP)的权重。其中 $W_0 \in \mathbf{R}^{\epsilon/r \times \epsilon}, W_1 \in \mathbf{R}^{\epsilon \times \epsilon/r}$ 。

空间注意力机制模块将经过通道注意力处理的 碎屑特征图作为新的输入特征图,首先对碎屑特征 图进行 GMP 和 GAP 操作,得到两个 $H \times W \times 1$ 的 碎屑特征图,分别记为 F_{max}^{s} 和 F_{avg}^{s} ,然后将 F_{max}^{s} 和 F_{avg}^{s} 两个碎屑特征图融合,经过一个 7 × 7 卷积核和 sigmoid 函数得到权重系数 $M_{s} \in R^{H \times W}$,最后 M_{s} 与 该模块的输入相乘生成最终特征。其具体如下式:

$$M_{s}(\mathbf{F}) = \sigma(f^{7\times7}([AvgPool(\mathbf{F}); MaxPool(F)])) = \sigma(f^{7\times7}([\mathbf{F}_{avg}^{s}; \mathbf{F}_{max}^{s}])), \qquad (2)$$

式中, **F** 指空间注意力机制输入的特征图, f^{xx7} 指滤 波器为 7 × 7 的卷积, F_{avg}^{s} 和 F_{max}^{s} 指经过最大池化和 平均池化的特征图。

2.2 DenseNet的ASPP

铣床碎屑特征提取编码流程中进行了多种卷积 与池化操作,同时 DeepLabV3+的 ASPP 模块分别 以采样率为6、12、18的3×3的卷积核对碎屑图像采 样,编码模块的采样和空洞卷积使得特征信息中所 含碎屑的像素信息减少,导致被用来计算的采样像 素数量更加稀疏,这种稀疏现象在二维情况下表现 更差。为获得更大的感受野及更密集的像素采样, 本文改进 DeepLabV3+的 ASPP 模块,以 DenseNet 的方式共享信息,不同扩张率的卷积核相互影响,实 现上一层卷积输出与下一层卷积输入特征图相连 接,最终将二者卷积的结果作为碎屑特征图在该卷 积层的输入。改进后的 ASPP 在特征提取过程中, 除了采样率为6的第一层卷积层外,其余两层都进 行了碎屑特征图多尺度融合,经过3层级联堆叠的 DebseNet 方式,增大了感受野并获得了更为密集的 像素采样方式^[15],改进的 ASPP 模块所提供的感受 野如图3所示。

设 H^{*k*} 表示空洞速率为*r* 且卷积核大小为*k* 的卷 积操作,则 ASPP 模块可如下式所示:

$$y = H_3^6(x) + H_3^{12}(x) + H_3^{18}(x)$$
(3)

空洞卷积的引入是为了在图像分割时保持特征 图像分辨率不变的情况下扩大感受野。对于输出信 号 y 和输入信号 x 关于空洞卷积的计算式如下:

$$y(i) = \sum_{k=1}^{K} x(i+r \cdot k) w(k) , \qquad (4)$$

式中,r表示空洞速率,w(k)表示滤波器在第 k 个 位置的参数,K 表示滤波器尺寸。空洞卷积是在卷 积核两个值之间插入r-1个零,从而增大感受野,并 且r值与感受野之间为正相关关系。对于一个卷积 核大小为 k 且空洞速率为r 的空洞卷积,其所能提供 的感受野大小如下式:

$$\mathbf{R} = (r-1) \times (k-1) + k \, . \tag{5}$$

对于 3 层级联空洞卷积层 R_1 、 R_2 和 R_3 能提供的感受野大小见下式:

$$R = R_1 + R_2 + R_3 - 2$$
 (6)





2.3 自适应空间特征融合

在 DeepLabV3+网络结构中,不同尺度碎屑特 征图都包含了不同级别的语义和细节信息,原 Deep-LabV3+的解码模块中仅使用输入流 1/4 大小的底 层特征图与编码模块的深层特征图进行融合,由于 碎屑边缘非常不规则,仅通过单层碎屑特征融合没 有考虑到其他层的有效信息,导致碎屑分割精度降 低。为了利用编码阶段产生的多尺度碎屑特征图, 本文引入 ASFF 网络,在解码阶段根据不同尺度的 特征自适应分配权重系数,实现多尺度高效融合,降 低细节方面的损失,提高分割精度,自适应特征融合 机制结构如图 4 所示。





本文选取 DeepLabv3+编码模块的 1/2、1/4 和 1/8 碎屑特征图进行融合,其中 1/4 的碎屑特征图和 通道数作为融合标准。如图 1 所示,对于 1/2 的碎屑 特征图,使用 1×1 的卷积调整通道数使其与 1/4 的 通道数一致,然后执行步长为 2、大小为 3×3 的卷积 操作,实现 2 倍的下采样;对于 1/8 的碎屑特征图,同 样先进行 1×1 的卷积操作,然后进行 2 倍的上采样, 得到统一后的特征图,根据 ASFF 对其进行多尺度 的特征融合。

在 ASFF 中, X_{ij}^1 , X_{ij}^2 和 X_{ij}^3 分别代表编码阶段 不同深度的特征图, a_{ij} , b_{ij} , c_{ij} 表示不同特征图中 特征点 (i,j) 的权重参数。其式如下:

$$a_{ij}\boldsymbol{X}_{ij}^{1} + b_{ij}\boldsymbol{X}_{ij}^{2} + c_{ij}\boldsymbol{X}_{ij}^{3} = \boldsymbol{Y}_{ij} , \qquad (7)$$

$$a_{ij} + b_{ij} + c_{ij} = 1$$
, (8)

式中, Y_{ij} 为融合后的新特征图,其中 a_{ij} 、 b_{ij} 、 $c_{ij} \in [0,1]$ 。

ASFF 为了实现权重参数 a_{ij} 、 b_{ij} 、 c_{ij} 的自适应 调整,使用误差梯度反向传播调节融合比例。通过 对统一尺寸后的 X_{ij}^1 、 X_{ij}^2 和 X_{ij}^3 特征图执行卷积核为 1×1的卷积操作,得到 A_{ij} 、 B_{ij} 、 C_{ij} 。最后使用 softmax 函数对 A_{ij} 、 B_{ij} 、 C_{ij} 进行处理得到权重参数 a_{ij} 、 b_{ij} 、 c_{ij} ,其式如下:

$$a_{ij} = \frac{\mathrm{e}^{\mathbf{A}_{ij}}}{\mathrm{e}^{\mathbf{A}_{ij}} + \mathrm{e}^{\mathbf{B}_{ij}} + \mathrm{e}^{\mathbf{C}_{ij}}}, \qquad (9)$$

$$b_{ij} = \frac{\mathrm{e}^{\boldsymbol{B}_{ij}}}{\mathrm{e}^{\boldsymbol{A}_{ij}} + \mathrm{e}^{\boldsymbol{B}_{ij}} + \mathrm{e}^{\boldsymbol{C}_{ij}}}, \qquad (10)$$

$$c_{ij} = \frac{\mathrm{e}^{\mathbf{c}_{ij}}}{\mathrm{e}^{\mathbf{A}_{ij}} + \mathrm{e}^{\mathbf{B}_{ij}} + \mathrm{e}^{\mathbf{C}_{ij}}} \,\,. \tag{11}$$

3 实验与分析

3.1 实验数据集及实验参数配置

本实验采用铣床加工中所产生的金属铅碎屑, 为保证实验数据的多样性,收集数控铣床在不同强 度光照下的金属碎屑。相机为 Basler 公司生产的 CMOS 工业面阵相机,型号为 acA640-300 gm,图片 分辨率为 640 pixel×480 pixel。使用图像标注工具 Labelme 对铣床碎屑样本集进行标注工作,为防止碎 屑图像训练过程中出现过拟合的现象,对实验铣床 碎屑数据集实现水平镜像、90°翻转等操作进行数据 的增广处理,最终获得训练数据集 7 000 张,测试集 840 张,验证集 1 000 张。铣床碎屑部分样本集图像 如图 5 所示。

为加快网络模型训练速度并降低损失,本实验 采用 Adam optimizer 函数作为参数优化器,训练初 始网络学习率为 0.007,每次送入网络训练的铣床碎 屑图像批次为4,训练的最大迭代次数为100。实验 平台基于 TensorFlow 深度学习框架下使用 python 3.6编程,在 CPU 为 i7-7800X、64 GB 内存、Ubuntu18.04 系统下使用 NVIDIA GeForce GTX 2080Ti 显卡、Cuda10.0进行训练加速。



图 5 部分碎屑图像 Fig. 5 Partial debris image

3.2 实验评价指标

本文选取深度学习分割领域中常见的 MIOU、 召回率(Recall)、精确率(Precision)以及 F_1 值(F_1 measure)作为实验图像分割的性能评判标准。

$$MIOU = \frac{1}{K+1} \sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij} + \sum_{j=0}^{k} p_{ji} - p_{ii}},$$
(12)

式中,共有 K + 1 个类别, p_{ii} 为预测类 i 正确的像素 数量, p_{ij} 为被预测为 i 类的像素总量, p_{ji} 为被预测 为 j 类的像素总量。

$$Precision = \frac{TP}{TP + FP} , \qquad (13)$$

式中, Precision 代表碎屑图中预测像素占总像素的 比例,即像素精度, TP 代表真实标签为碎屑且被分 割为碎屑的图像, FP 代表真实标签为背景但被分割 为碎屑的图像。

$$Recall = \frac{TP}{TP + FN} , \qquad (14)$$

式中, FN 表示真实标签为铣床碎屑, 但被分割为背景的像素个数。

$$F_1 = \frac{2PA \times Recall}{Precision + Recall} , \qquad (15)$$

式中, F₁值为召回率和平均精度的加权调和平均值,可以对算法的分割性能做综合评价。

3.3 模型训练结果

本文实验网络训练结果如图 6 所示,图中示出 DeepLabv3+与本文方法的损失值随迭代次数的变 化过程。二者损失值都随迭代次数逐渐增加而趋于 稳定,本文方法收敛速度更快,CBAM 注意力机制有 效结合了铣床碎屑图像的空间和通道维度,并结合 DenseNet 的 ASPP 加强了特征传播,减少冗余碎屑 图像特征的计算量。当实验训练损失值下降到平缓 状态后,模型基本完成收敛,选取本次训练的最低权 重作为最终测试模型,进行本次实验性能的检验。





3.4 消融实验

为了验证将 CBAM 注意力机制嵌入 Xception 网络中的有效性,以及在编码阶段采用 DenseNet 的 ASPP 模块和解码阶段采用 ASFF 模块的必要性,通 过添加单个模块和组合多个模块,设计了 4 组消融 实验并根据多组评价指标数据展示其对网络的性能 提升,实验结果如表 1 所示。

表 1 改进型 DeepLabV3+消融实验结果 Tab. 1 Results of DeepLabV3+ ablation experiments

Test	CBAM	DenseNet	ASFF	MIOU
1	—	_	_	0.819
2	\checkmark	—	—	0.822
3	\checkmark	\checkmark	—	0.826
4	\checkmark	\checkmark	\checkmark	0.839

实验1为原 DeepLabV3+网络,实验2、3和4 分别为在实验1的基础上逐个添加CBAM模块、 DenseNet模块、ASFF模块。由实验1、2可知嵌入 了 CBAM 注意力机制后碎屑分割模型的 MIOU 提 升了 0.003,该方法说明 CBAM 注意力机制通过优 化通道与空间权重使得碎屑的位置以及轮廓特征提 取的更加全面,提升分割精度。实验3在加入 CBAM 模块的基础上使用 DenseNet 的 ASPP 模块, 模型的 MIOU 提升了 0.004,该方法增强各支路间 碎屑特征图的相关性,增大模型感受野,从而提升预 测模型的整体效率。实验4同时添加 CBAM 模块、 DenseNet 模块与 ASFF 模块, MIOU 达到了 0.839, 该实验证明在解码过程中采用 ASFF 模块提取了图 像高层特征的语义信息与底层的细粒度特征,增强 解码能力,使得分割效率进一步提升。这表明实验 中对 DeepLabV3+算法的改进策略均能提升算法对 碎屑目标的分割效果。图 7 为表 1 中各组实验的可 视化结果。

为研究 CBAM 注意力模块在碎屑特征提取的作用,采用了梯度类激活映射方法对碎屑特征提取部分进行可视化,如图 8(c)、(d)所示,颜色较深碎屑区域的变化会极大影响分割的预测结果,颜色越深对预测结果的权重贡献度越大。由图 8 可知,引入 CBAM 注意力机制后,图中碎屑部分聚集了更多红 色区域且颜色更深,表明 CBAM 注意力机制使得网 络集中于碎屑的特征,忽略背景等无关特征的影响。

3.5 分割性能分析

为了验证算法的鲁棒性,将车间的光照强度分为偏暗光200lx、正常光600lx、偏强光1000lx和 1600lx超强光4个等级,对4种光照强度的碎屑图 像进行实验。如图9所示,当光照强度为200lx时, 碎屑与背景对比度较低;当光照强度为600lx时,目 标特征较为清晰;当光照强度为1000—1600lx时, 表现出碎屑整体或者局部反光现象。在下列4组图



图 7 消融实验可视化结果:(a) 原图;(b) 实验 1;(c) 实验 2;(d) 实验 3;(e) 实验 4 Fig. 7 Visualization results of ablation experiments: (a) Original image; (b) Test 1; (c) Test 2;(d) Test 3;(e) Test 4



图 8 可视化效果对比图:(a) 原图; (b) 标注图; (c) DeepLabV3+; (d) CBAM-DeepLabV3+ Fig. 8 Comparison of visual effects:(a) Original image; (b) Label image; (c) DeepLabV3+; (d) CBAM-DeepLabV3+



图 9 不同光照强度下分割结果对比图:(a) 偏暗光照分割结果;(b) 正常光照分割结果; (c) 偏强光照分割结果;(d) 超强光照分割结果

Fig. 9 Comparison of segmentation results under different light intensities: (a) Weak light segmentation results;

(b) Normal light segmentation results; (c) Segmentation results of partial strong light; (d) Highlight light segmentation results



图 10 不同分割网络预测对比图:(a) 单目标分割结果;(b) 多目标分割结果;(c) 堆叠多目标分割结果 Fig. 10 Comparison chart of predictions of different segmentation networks:(a) Single object segmentation results; (b) Multi-objective segmentation results; (c) Stack multi-object segmentation results

像中,本文算法对碎屑图像均可以进行准确的分割, 避免光照强度变化引起的误分割问题,反映了该算 法具有良好的鲁棒性。

为进一步验证本文方法的性能,将3组正常光 照强度下的不同类型碎屑图像输入训练好的多种模 型中,由图 10 可知本文方法的分割效果明显优于其 他网络,本文方法提取的碎屑轮廓更加清晰,且提取 金属碎屑较为完整且连续,具有良好鲁棒性,而其他 3 种金属碎屑的分割方法均存在图像断连和细节部 分较为模糊的问题。其中,PSPNet 对于碎屑提取效 果最差,不能完全体现碎屑几何形状及对部分较小碎屑出现的漏分割现象;对于 U-Net,虽然可以看出碎屑大体轮廓,但是出现较多的断连碎屑,导致较多孤立点的存在,不能完全体现碎屑的基本特征,且对于堆叠状态下的图像分割效果不佳;对于 Deep-LabV3+算法,虽然可以看出碎屑的整体结构,但在部分碎屑末端仍然存在断连以及碎屑边缘不平滑现象。本文改进后的 DeepLabV3+算法由于嵌入了注意力机制模块,并在 ASSP 模块中采用 DenseNet 的方式,极大地缓解了采样过程中特征图信息的丢失问题,使边缘识别更加清晰准确,表现出良好的分割性能。

表 2 为测试图像客观评价指标比较,明显看出 PSPNet、U-Net、DeepLabV3+模型与本文提出的方 法在平均精度、召回率、 F_1 值和 MIOU 4 个方面的表 现都有较大差距,本文方法的像素精度达到 0.871, 召 回 率 达 到 0.891, F_1 值 达 到 0.881, MIOU 达 0.839。由表 2 可得, PSPNet 分割性能较差,相比其 他 3 种模型,本文算法在铣床碎屑分割性能上更具 优势。

表 2 不同分割网络模型性能比较

 Tab. 2
 Performance comparison of different segmentation network models

Models	MIOU	Precision	Recall	${F}_1$
PSPNet	0.771	0.826	0.899	0.860
U-Net	0.789	0.812	0.898	0.852
DeepLabV3+	0.819	0.845	0.894	0.868
Ours	0.839	0.871	0.891	0.881

4 结 论

本文针对铣床工业加工中碎屑轮廓多样且车间 光照不均,导致分割效果较差的问题,提出了一种改 进 DeepLabV3+铣床碎屑分割方法。首先,在 Deep-LabV3+编码端 Xception 嵌入注意力机制模块,利 用通道和位置像素间的特征信息提升有用特征表达 权重,使网络更为精准地分割目标碎屑边缘。其次, 构建了 DenseNet 的 ASPP 模块,增加了感受野的同 时丰富了特征尺度多样性,提升有用信息利用率。 最后,通过分配多尺度碎屑特征图的融合权重,实现 解码过程中碎屑特征信息的高效运用,从而提升网 络分割性能。本文与其他分割算法进行对比,鲁棒 性与分割性能均有较大提高。

在后续研究中,结合铣床碎屑图像的分割效率

与在铣床上部署实时监测的需求,在保证网络分割 精度的情况下,精简模型并尽可能研究多种不同材 质的碎屑图像分割是关于铣床碎屑图像分割下一步 需要研究的方向。

参考文献:

- [1] GHOSH S, DAS N, DAS I, et al. Understanding deep learning techniques for image segmentation [J]. ACM Computing Surveys, 2019, 52(4):1-35.
- [2] LIU X, DENG Z, YANG Y. Recent progress in semantic iMage segmentation[J]. Artificial Intelligence Review, 2019, 52(2):1089-1106.
- [3] YU H, YANG Z, TAN L, et al. Methods and datasets on semantic segmentation: A review [J]. Neurocomputing, 2018, 304:82-103.
- [4] MINAEE S, BOYKOV Y, PORIKLI F, et al. Image segmentation using deep learning: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44 (7):3523-3542.
- [5] ZHANG J, LU C, LI X, et al. A full convolutional network based on DenseNet for remote sensing scene classification
 [J]. Mathematical Biosciences and Engineering, 2019, 16
 (5):3345-3367.
- [6] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network [C]//IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York; IEEE, 2017;6230-6239.
- [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, 2017, 40(4): 834-848.
- [8] BAHETI B, INNANI S, GAJRE S, et al. Semantic scene segmentation in unstructured environment with modified DeepLabV3+[J]. Pattern Recognition Letters, 2020, 138: 223-229.
- [9] DENG J L, DONG L H, SONG W, et al. Processing of seabed polymetallic nodule images based on sea-thru and mask R-CNN [J]. Mining and Metallurgical Engineering, 2022,42(2):9-13.
- [10] YUAN L, DUAN Y. A method of ore image segmentation based. on deep learning[C]//International Conference on Intelligent Computing, August 15-18, 2018, Wuhan, China.

Cham: Springer, 2018: 508-519.

- [11] LIU C, CHEN L C, SCHROFF F, et al. Auto-deeplab: Hierarchical neural architecture search for semantic image segmentation [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019:82-92.
- [12] KASSANI S H, KASSANI P H, KHAZAEINEZHAD R, et al. Diabetic retinopathy classification using a modified xception architecture [C]//2019 IEEE International Symposium on Signal Processing and Information Technology (IS-SPIT), December 10-12, 2019, Ajman, United Arab Emirates. New York: IEEE, 2019:1-6.
- [13] TANG J, GONG C, GUO F, et al. Geo-localization based on CNN feature matching[J]. Optoelectronics Letters, 2022, 18(5): 300-306.

- [14] LI D, HU X, ZHANG H, et al. A GAN based method for multiple prohibited items synthesis of X-ray security image[J]. Optoelectronics Letters, 2021, 17(2):112-117.
- [15] BADRINARAYANAN V,KENDALL A,CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 (12): 2481-2495.

作者简介:

刘秀平 (1981-),男,博士,副教授,硕士导师,主要从事机器视觉、 图像处理方面的研究.