基于轻量级 MobileNetV2-DeeplabV3+的棒材分割方法

汤维杰,方 挺,韩家明,袁东祥
安徽工业大学 电气与信息工程学院,安徽 马鞍山 243000

摘 要:针对当前语义分割模型为提升像素分割精度,不断增加算法复杂度,导致模型出现参数量大,耗时长,难以部署至工业现场等问题,提出一种基于轻量级 MobileNetV2-DeeplabV3+模型的棒材分割算法。算法为平衡像素分割精度、模型参数量和算法检测速度,在原网络基础上做出一系列改进:将原有的 Xception 主干网络替换为轻量级 MobileNetV2 网络以降低模型参数量与计算复杂度;在空洞空间金字塔池化(ASPP)模块基础上密集连接各空洞卷积以获得更大的感受野,更加密集的像素采样,并扩大输出特征覆盖的语义信息;使用深度可分离卷积(DSConv)替代 ASPP 模块中的标准卷积进一步降低模型的计算复杂度;此外,引入有效通道注意力(ECA)模块聚焦目标边缘特征,增强特征图通道信息提取的效果。实验表明:改进后的模型在棒材数据集下平均交并比(MIOU)为 89.37%,平均像素精度(MPA)为 94.57%,帧率(FPS)为 33.09 帧/s,模型参数量为 33.6 M。与 U-net、M-PSPNet、M-DeeplabV3+等模型相比,改进后算法的 MIOU 值与 MPA 值略低于最佳值,但仍处于较高水准,模型参数量小,FPS 值得到较大提升。实验表明:改进后的算法能较好地平衡分割精度和算法实时性,能满足部署至工业现场的需求。

关键词:语义分割;DeepLabv3+模型;轻量级;棒材
中图分类号:TF34;TP391.41 **文献标识码**:A doi:10.16055/j.issn.1672-058X.2024.0003.009

Bar Segmentation Method Based on Lightweight MobileNetV2-DeeplabV3+

TANG Weijie, FANG Ting, HAN Jiaming, YUAN Dongxiang
School of Electrical and Information Engineering, Anhui University of Technology, Anhui Maanshan 243000, China

Abstract: In order to improve the pixel segmentation accuracy of the current semantic segmentation model, the algorithm complexity continues to increase, resulting in a large number of parameters, time-consuming, and difficulty in deploying to industrial sites. A bar segmentation algorithm based on the lightweight MobileNetV2-DeeplabV3+ model was proposed. The algorithm made a series of improvements based on the original network in order to balance the pixel segmentation accuracy, the number of model parameters and the detection speed of the algorithm. The original Xception backbone network was replaced with a lightweight MobileNetV2 network to reduce the number of model parameters and computational complexity. On the basis of the Atrous Spatial Pyramid Pooling (ASPP) module, the atrous convolutions were densely connected to obtain a larger receptive field and denser pixel sampling, and to enlarge the semantic information covered by the output features. The computational complexity of the model was further reduced by using deep separable convolution (DSConv) instead of the standard convolution in the ASPP module. In addition, an effective channel attention (ECA) module was introduced to focus on the target edge features and enhance the effect of channel information extraction in the feature maps. The experiment showed that the improved model achieved a mean intersection

收稿日期:2021-03-05 **修回日期**:2021-05-18 **文章编号**:1672-058X(2024)03-0066-06
基金项目:安徽工业大学校青年基金(QZ202109);安徽工业大学校青年基金(QZ202109).
作者简介:汤维杰(1997—),男,安徽马鞍山人,硕士研究生,从事计算机视觉研究.
通讯作者:方挺(1975—),男,安徽马鞍山人,教授,博士,从事计算机视觉研究. Email:flyting_69@163.com.
引用格式:汤维杰,方挺,韩家明,等. 基于轻量级 MobileNetV2-DeeplabV3+的棒材分割方法[J]. 重庆工商大学学报(自然科学版),2024,41(3):66—71.
TANG Weijie, FANG Ting, HAN Jiaming, et al. Bar segmentation method based on lightweight MobileNetV2-DeeplabV3+[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2024, 41(3): 66—71.

over Union (MIOU) of 89.37%, a mean pixel accuracy (MPA) of 94.57%, a frame rate of 33.09 frames per second (FPS), and a model parameter size of 33.6 M on the bar dataset. Compared with the models of U-net, MPSPNet, and M-DeeplabV3+, the MIOU and MPA values of the improved algorithm were slightly lower than the best values, but still at a high level, with a small number of model parameters and a significant increase in FPS value. The example shows that the improved algorithm can better balance the segmentation accuracy and the real-time performance of the algorithm, and can meet the needs of deployment to industrial sites.

Keywords: semantic segmentation; DeepLabv3+ model; lightweight; bar

1 引　言

棒材作为工业生产的基础原材料,广泛用于建筑、机械、汽车、船舶等工业领域。棒材尺寸测量是棒材生产过程中的一道重要工序,棒材尺寸的合格率是衡量一个钢铁企业好坏的重要依据之一[1],而棒材尺寸测量方法中最关键的技术为棒材边缘分割技术。

当前基于图像处理的边缘分割方法主要分为传统图像处理方法和深度学习方法。传统图像处理方法一类是利用图像梯度的一阶或二阶导数的相关特点来检测图像边缘,常见的微分算子有 Sobel 算子[2],Prewitt 算子[3]和 Canny 算子[4]等。这些梯度算法的优点是效率高,但易造成边缘模糊,抗干扰能力差。另一类为基于强度,梯度和纹理等人工设计特征的学习方法。Martin 等[5]将纹理、亮度、颜色特征信息组合成 Pb 特征并输入至分类器生成边缘线条。该方法相较于早期的图像梯度算法,边缘检测效果得到较大提升,但其步骤多,成本高。刘建培等[6]为解决简单背景下圆钢截段长度尺寸测量问题,先采用高斯滤波、灰度化、二值化等方法预处理圆钢图像,再采用最大二值面积特征确定圆钢的最小外接矩形,进而测得长度尺寸。该算法简单易行,实时性高,但不适用于复杂背景下的目标分割问题。倪超等[7]针对复杂场景下粘连棒材难以进行图像分割的问题,提出强弱粘连两步标记的分水岭算法,该方法采用一种新型棒材中心位置标记方法,根据强弱相连特征生成棒材的初始标记再使用分割算法分割棒材,从而实现棒材的准确计数。该算法由于采用两步法分割,多种算法融合,分割效果优良。但由于算法过于复杂,导致模型的实时性差,泛化能力弱。

上述基于人工设计的特征在特定场景检测效果良好,但设计复杂,特征寻找困难且缺少语义信息,故而易受环境影响。基于深度学习的分割方法不但具有强大的特征表达能力,更能提取传统方法所没有的高层次语义信息,这为边缘检测技术带来了新的突破。Long 等[8]利用现有分类网络 AlexNet[9]、VGG[10]和 GoogLeNet[11]构建 FCN(Fully Convolutional Network)网络,并通过微调将网络应用到分割任务中,实现了图像分割问题端对端训练的突破,但 FCN 网络训练繁琐,需要训练 3 次且恢复空间信息时上采样仅采用反卷积操作,上采样像素过于稀疏,输出图像不够精细。Ronneberger 等[12]提出新型分割模型 U-Net,针对医学图像数据集难以获取且数量稀少的问题,在 FCN 网络的基础上进行改进,仅需要少量图像即可获得较为精准的分割效果。模型构建简单,只需训练一次,同时将上采样方法改进为采用跳层连接将浅层特征信息与高层语义信息相融合,但由于模型结构单一,未能考虑像素间的联系。Zhao 等[13]研究 FCN 网络,认为该网络缺乏对全局语境信息的收集,从而提出 PSPNet 分割模型,PSPNet 通过池化操作聚合不同区域的上下文来获取全局上下文信息,增强网络的特征提取能力。该算法虽然分割精度较高,但模型参数量与计算复杂度均十分巨大。Chen 等[14]提出了一种使用空洞卷积的语义分割模型 DeeplabV3+,该模型引入采用级联不同空洞卷积率的空洞卷积来获取多尺度上下文信息的空洞空间金字塔(ASPP)模块,该模块避免了因重复池化和下采样带来的位置信息丢失,并添加解码器模块逐渐恢复空间信息以获取更清晰的对象边界,在一定程度上减少了模型处理时间。左纯子等[15]通过将全局上采样与 DeeplabV3+网络相结合,提出基于深度神经网络的煤尘图像分割方法,该方法通过重新设计一组针对小物体分割的空洞率,使得小型煤尘颗粒得到准确分割,但算法耗时长,模型复杂,对硬件要求较高。汝承印等[16]针对图像分割模型难以搭载无人机等嵌入式设备检测高压电线上的绝缘子缺陷问题,提出采用 MobileNet-SSD 算法和 MobileNetV2-DeeplabV3+算法相结合的方式识别绝缘子缺陷,首先使用 MobileNet-SSD 算法对绝缘子准确定位并分类,再使用 MobileNetV2-DeeplabV3+算法对绝缘子缺陷分割。该方法通过使用轻量级主干网络有效减小了模型复杂度,但不可避免地带来分割精度的下降。此外,先采用目标检测算法再使用分割模型并不能为分割精度带来明显的提升,若无特殊分类需求,无须使用。

这些基于深度学习的语义分割模型为提升像素分割精度,不断增加算法复杂度,导致模型出现参数量大,耗时长,难以部署至工业现场等问题。基于上述分析,提出一种基于轻量级 MobileNetV2-DeeplabV3+模

型的棒材分割算法。该算法以轻量级网络 MobileNetV2 为主干网络,有效降低模型复杂度;改进了 ASPP 模块,扩大感受野的同时捕获更加密集的像素特征;引入深度可分离卷积和有效通道注意力机制降低计算复杂度并增强网络的特征提取能力。

2 改进的 DeepLabv3+网络

2.1 DeepLabv3+整体框架

考虑 DeeplabV3 网络没有包含足够丰富的浅层信息,导致提取的物体边界,尤其是尖锐物体边缘存在不够清晰等问题,DeeplabV3+网络在 DeeplabV3 架构的基础上引入编码器-解码器(Encoder-Decoder)[17]结构替代原有的概率图模型(CRF)结构,该结构中的解码器部分可以修复分割目标的边缘信息,并减少部分处理时间。编码器中主要采用以 Xception 为主干网络提取目标特征,通过使用 ASPP 模块中的多个空洞卷积获取高层信息,提高模型感受野;解码器中先将从 ASPP 模块中获取的高层语义信息与从 Xception 主干网络中捕获的浅层信息经 1×1 卷积通道压缩后相融合,再通过两个 3×3 卷积逐步恢复空间信息以获得更加细化的目标边缘。

2.2 改进的 DeepLabv3+整体框架

改进的 DeeplabV3+模型仍由编码器和解码器两部分构成,模型结构如图 1 所示。编码器部分使用 MobileNetV2 替换 Xception 作为新的主干网络,提取待分割目标的特征信息;通过改进 ASPP 模块,使用含有不同卷积率的密集连接空洞卷积(DenseASPP)提取高层语义信息;并将 ASPP 模块中使用的标准卷积替换为深度可分离卷积减少模型参数量,降低模型复杂度。在解码器中,浅层特征信息和高层语义信息先采用 ECA 注意力机制聚焦目标特征,提高分割效果。高层语义特征经过 4 倍双线性插值上采样后与通道压缩后的低层特征融合,通过 3×3 卷积和 4 倍双线性插值上采样恢复至原图大小并获得精细的目标边缘。

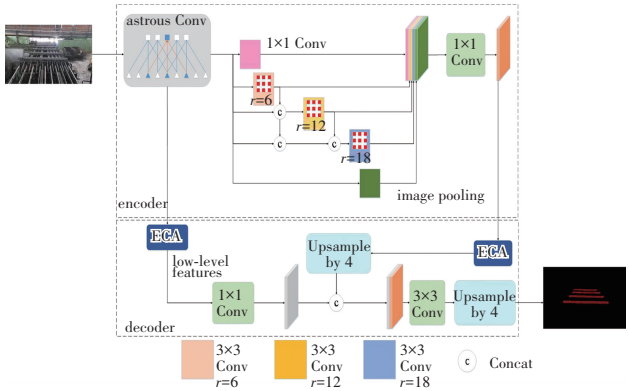


图 1 改进的 DeepLabv3+模型结构

Fig. 1 Improved DeepLabv3+ model structure

2.3 密集连接空洞空间金字塔池模块

在神经网络中,常使用多次池化操作提高模型的感受野,但池化操作在增强感受野的同时也会带来空间分辨率降低的问题,这将导致特征信息和平移不变性的丢失。为解决特征图分辨率与感受野大小之间的矛盾,DeeplabV3+网络中引入 ASPP 模块,通过使用多个不同膨胀率的空洞卷积捕获多尺度特征图信息。为了获取足够大的感受野就需使用足够大的膨胀率,但随着膨胀率的增加,膨胀卷积会变得越来越无效且逐渐失去建模能力[17]。此外,ASPP 模块中采用不同膨胀率的空洞卷积采集到的特征信息像素稀疏,未充分利用图像信息。基于上述问题,本文采用密集连接各空洞卷积(DenseASPP)[18]来获得更加密集的像素采样,膨胀率为 6、12、18 的空洞卷积互相关联,低膨胀率空洞卷积的输出与主干网络的输出先做 Concat 操作后再输入至下一层中,使得两个高膨胀率的空洞卷积中都涵盖了特征图的多尺度信息,并通过堆叠 3 层空洞卷积扩大感受野的范围,DenseASPP 模块结构图如图 1 所示。

等效感受野的尺寸计算如式(1)所示:

$$R=(d-1)\times(K-1)+K \tag{1}$$

其中,R 为感受野尺寸,d 为膨胀率,K 为卷积核尺寸。

堆叠后的感受野尺寸计算如式(2)所示:

$$R=R_1+R_2+\cdots R_n-(n-1) \tag{2}$$

其中,R 为堆叠后的感受野尺寸,$R_1,R_2,\cdots R_n$ 为各空洞卷积感受野尺寸。

以膨胀率为 6、12、18 的 3×3 空洞卷积为例,ASPP 模块中的等效感受野尺寸仅需计算膨胀率为 18 的空洞卷积的感受野,大小为 37。DenseASPP 模块中则需叠加三个空洞卷积的感受野,大小为 73,感受野大小与原模块相比扩大明显。

2.4 深度可分离卷积

深度可分离卷积[19]主要分为两个过程:逐通道卷积(Depthwise Convolution)和逐点卷积(Pointwise Convolution),图 2 为深度可分离卷积与标准卷积对比图。逐通道卷积中卷积核个数与通道数相等,一个通道只被一个卷积核卷积,最终得到的通道数与输入的通道数相一致。逐点卷积与标准卷积类似,其卷积核大小为 1×1。标准卷积与深度可分离卷积的参数量比值和计算量比值如式(3)和式(4)所示:

$$\frac{P_c}{P_d}=\frac{C_i\times k^2+C_i}{C_i\times k^2\times C_n+C_n} \tag{3}$$

$$\frac{C_c}{C_d}=\frac{C_i\times H\times W\times(k^2+C_n)}{C_i\times H\times W\times k^2\times C_n} \tag{4}$$

其中,P_c 和 P_d 分别为标准卷积和深度可分离卷积的参

数量；C_c 和 C_d 分别代表标准卷积和深度可分离卷积的计算量；C_i 和 C_n 分别为输入通道数和输出通道数；H 和 W 分别表示输入图像的长和宽；k 为卷积核大小。

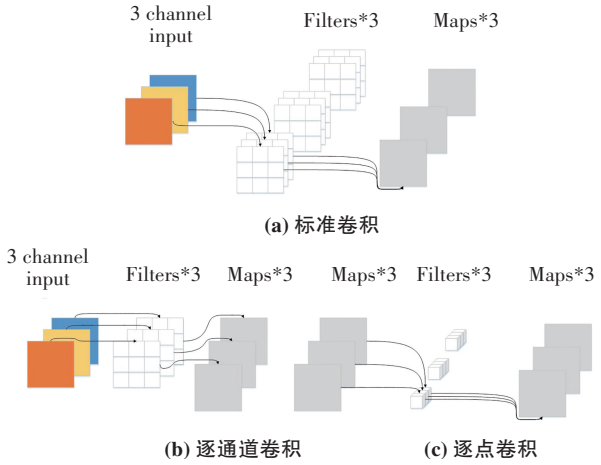


图 2　深度可分离卷积与标准卷积对比

Fig. 2　Comparison of depthwise separable convolution and standard convolution

由式(3)可近似得出深度可分离卷积的计算量约为标准卷积的 $1/C_n$。当输出通道数较大时，使用深度可分离卷积的计算量约为标准卷积的 $1/k$，以标准卷积核 3×3，输出通道数 256 为例，其参数量约为标准卷积的 1/256，计算量约为标准卷积的 1/9，DenseASPP 模块的参数量和计算复杂度显著减小。

2.5　ECA 注意力机制

近年来，SENet 等[20]通道注意力机制广泛应用于 CNN(Convolutional Neural Network)中且表现良好。考虑现有方法大多倾向于增加算法复杂度改进 SENet 模型以获得更好的性能，本文模型引入一种较为轻量化的 ECA(Efficient Channel Attention)[21]模块，在提高网络性能的同时保持较低的计算复杂度，SENet 与 ECA 模块对比如图 3 所示。

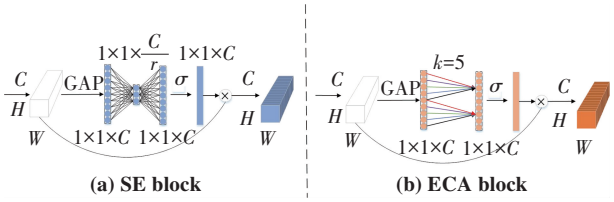


图 3　SENet 与 ECA 模块对比

Fig. 3　Comparison of SENet and ECA modules

两模块均使用全局平均池(GAP)获得聚合特征，获得特征后 SENet 先进行通道压缩，再使用两个捕获非线性跨通道交互的 FC(Fully Connected Layer)层和一个 s 形函数生成通道权重。Wang 等[21]经研究发现通道压缩会影响通道注意预测效果，采用所有通道的依赖是耗时久且非必要的。ECA 为避免通道压缩，通过执行大小为 k 的快速 1D 卷积和 Sigmod 激活函数来捕获本地跨通道交互并生成通道权重 ω，ω 的计算公式如式(5)所示：

$$\omega = \sigma\left(\sum_{j=1}^{k} \alpha^j y_i^j\right), y_i^j \in \Omega_i^k \tag{5}$$

其中，Ω_i^k 为 y_i^j 的 k 个邻域通道，σ 为 Sigmod 激活函数。

k 由通道维度 C 的映射自适应地确定。k 的计算如式(6)所示：

$$k=\psi(C)=\left|\frac{\log_2(C)}{\gamma}+\frac{b}{\gamma}\right| \tag{6}$$

其中，$\gamma=2$，$b=1$。

ECA 模型有效避免了因维度缩减带来的不良影响，且适当的跨通道交互可以在降低模型的复杂度的同时维持模型良好性能。如图 3 所示，本文在解码器中添加了 ECA 模块，在几乎不增加计算负担的同时聚焦棒材目标的边缘信息，有效增强网络的特征提取能力。

3　实验验证与分析

3.1　实验设置

实验使用的深度学习框架为 Pytorch1.7.0，编程语言为 Python3.8，硬件配置：CPU 为 E5-2650 v4，GPU 为 Tesla P40。模型使用 ImgaeNet 预训练模型初始化，采用随机梯度下降法训练，损失函数采为交叉熵损失函数。考虑本文所使用的数据集大小及硬件条件，采用冻结训练方式，在预训练模型参数上进行微调，冻结训练阶段 batchsize 为 8，lr 为 5e-4，解冻阶段 batchsize 为 4，lr 为 5e-5。

使用的图像分割数据集采自某钢铁厂热轧车间，图像大小为 1 280×720。由于工业现场灰尘大，环境较为恶劣，采集到的部分图像数据存在模糊不清现象，不能直接输入至网络中用以学习和提取棒材的特征。经预处理后数据集共有图像 275 张，通过对数据集使用随机旋转、翻转、缩放、裁剪等操作使数据集得到扩充。

设置的评价指标为平均交并比(MIOU)、平均像素准确率(MPA)、模型参数量和帧数(FPS)。平均交并比为网络对目标分割结果和真实值的交集与并集的比值，求和再求平均值；平均像素准确率为棒材被正确分类的像素数与真实值的像素数的比值；模型参数量用以代表模型的复杂度；帧数表示模型的检测速度。

3.2　损失函数

损失函数一般用于评估和监测神经网络模型的训练效果，棒材图像分割本质上是对棒材像素的分类问题，而交叉熵损失函数(Cross Entropy Loss，CE)常用于和 softmax 函数相结合来解决此种分类问题，softmax 函数可以将输出结果转化为[0，1]之间的概率分布。因

此本文损失函数采用交叉熵损失函数结合 softmax 函数,交叉熵函数值越小,表明两个类别离得越近。二分类时的交叉熵损失函数如式(7)所示:

$$L=-[y\cdot\log(p)+(1-y)\cdot\log(1-p)] \tag{7}$$

其中,y 为数据集中的分类标签,取值为 0 表示错误类别,为 1 表示正确类别;p 为分类预测正确的概率。

当样本为多分类时交叉熵损失函数如式(8)所示:

$$L=-\sum_{c=1}^{M}y_c\log(p_c) \tag{8}$$

其中,M 为数据集中总类别数;c 为对应的某一类别;y_c 取值 0 或 1,当前预测的类别与样本中标记的类别相同时取 1,否则取 0;p_c 为样本正确预测为 c 类别的概率。

3.3 实验对比分析

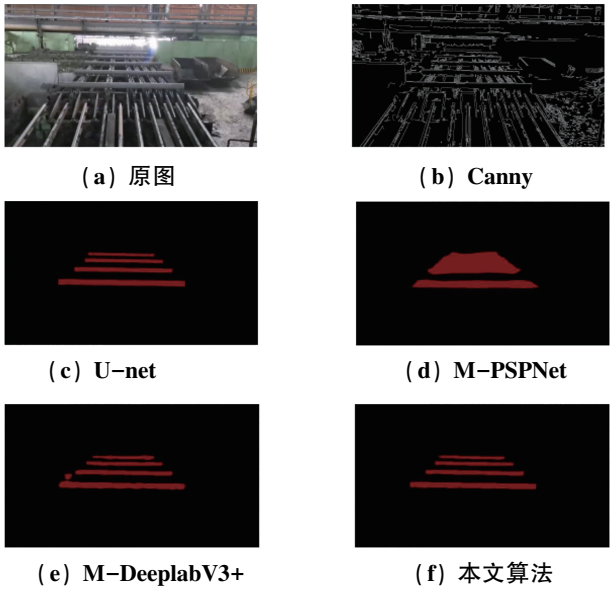
表 1 为模型与其他经典模型在棒材数据集上的结果对比,其中将 PSPNet 的主干网络替换为 MobileNetV2,记为 M-PSPNet。由表 1 中可以得出 U-net 的 MIOU 和 MPA 取得最佳效果,分别为 92.27%和 96.10%,但其参数量高达 94.9 M,模型过于复杂,且 FPS 仅有 7.01 帧/s;替换为轻量级主干网络的 M-PSPNet 的参数量仅为 9.30 M,帧数也最高,达到 70.91 帧/s,但其 MIOU,MPA 均为最低,实际分割效果差,难以满足分割精度要求;M-DeeplabV3+的 MIOU 和 MPA 略低于 U-net,但其参数量较小,FPS 也处于较高水准;在此基础上改进后的本文算法参数量略有提升,帧率有所增加,MIOU,MPA 分别提高了 2.83%和 0.08%,能在准确分割棒材目标的同时,保证算法实时性能够满足工业现场需求。

表 1 部分算法对比

Table 1 Comparison of some algorithms

算 法	MIOU/%	MPA/%	参数量/M	FPS/(帧·s^{-1})
U-net	92.27	96.10	94.9	7.01
M-PSPNet	79.61	90.81	9.30	70.91
M-DeeplabV3+	86.54	94.49	22.4	31.52
本文算法	89.37	94.57	33.6	33.09

图 4 为部分算法分割效果对比图,基于传统图像处理的 Canny 算法对于复杂背景下的棒材目标区分不明显,且出现棒材下边缘与背景相融合,未能分割等问题;U-net 中棒材目标分割清楚,平滑无毛刺;M-PSPNet 仅能分割出目标主体,各目标连成区域,无法精准分割出各目标边缘,不符合实际要求;M-DeeplabV3+能够较好地分割出目标边缘,但存在误检区域且部分边缘存在毛刺,不够平滑等现象;本文算法能准确分割出各目标边缘,有效减少误检与边缘毛刺现象,能较好地平衡分割精度与算法实时性。



(a) 原图 (b) Canny (c) U-net (d) M-PSPNet (e) M-DeeplabV3+ (f) 本文算法

图 4 部分算法分割效果对比

Fig. 4 Comparison of segmentation effects of some algorithms

表 2 为各模块改进效果对比,在 ASPP 模块中加入深度可分离卷积后,MIOU、MPA 略有降低,但 FPS 增长了 3.54 帧/s,模型的计算复杂度下降明显;当解码器中引入 ECA 注意力机制后 MIOU、MPA 分别提高了 1.43%、0.98%,帧率较原网络仍有较大提高;将 ASPP 模块改进为 DenseASPP 模块后 MIOU、MPA 得到较大提升,分别提高了 2.35%、1.1%,表明密集连接多个空洞卷积能提取到更多有用的高层特征信息;同时添加 ECA 模块和 DenseASPP 模块后 MIOU、MPA 分别提高了 2.83%、0.08%,FPS 也提高了 1.57 帧/s,在降低模型复杂度的同时带来分割效果的提升。

表 2 各模块改进效果对比

Table 2 Comparison of the improvement effects of each module

算 法	MIOU/%	MPA/%	FPS/(帧·s^{-1})
Base	86.54	94.49	31.52
Base+DS	86.15	93.73	35.06
Base+DS+ECA	87.97	95.47	34.36
Base+DS+DASPP	88.89	95.59	33.18
Base+DS+ECA+DASPP	89.37	94.57	33.09

4 结 论

针对当前语义分割模型大多通过增加模型复杂度提高像素分割精度,导致模型出现参数量大,耗时长,难以部署至工业现场等问题,提出一种基于轻量级 MobileNetV2-DeeplabV3+模型的棒材分割算法。该算法为平衡像素分割精度、模型参数量和算法检测速度,在原网络基础上做出一系列改进:将原有的 Xception 主干网络替换为轻量级 MobileNetV2 网络以降低模型参数量与计算复杂度;在 ASPP 模块基础上密集连接各空洞卷积以获得更大的感受野,更加密集的像素采样,

并扩大输出特征覆盖的语义信息;使用深度可分离卷积替代 ASPP 模块中的标准卷积进一步降低模型的计算复杂度;此外,引入 ECA 模块聚焦目标边缘特征,增强特征图通道信息提取的效果。实验表明改进后的模型在棒材数据集 MIOU 为 89.37%,MPA 为 94.57%,FPS 为 33.09 帧/s,模型参数量为 33.6 M。与 U-net、M-PSPNet、M-DeeplabV3+等模型相比,改进后算法的 MIOU 值与 MPA 值略低于最佳值,但仍处于较高水准,模型参数量小,FPS 值得到较大提升。实例表明:改进后的算法能较好地平衡分割精度和算法实时性,能满足部署至工业现场的需求。

参考文献(References):

[1] 侯幸林, 周培培, 赵景波, 等. 面向机器视觉的不锈钢棒材表面螺纹缺陷检测[J]. 重庆理工大学学报(自然科学), 2022, 36(5): 109—114.
HOU Xing-lin, ZHOU Pei-pei, ZHAO Jing-bo, et al. Machine vision-oriented thread defect detection of stainless steel bars[J]. Journal of Chongqing University of Technology (Natural Science), 2022, 36(5): 109—114.

[2] KITTLER J. On the accuracy of the Sobel edge detector[J]. Image and Vision Computing, 1983, 1(1): 37—42.

[3] TORRE V, POGGIO T A. On edge detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986(2): 147—163.

[4] CANNY J. A computational approach to edge detection[J]. IEEE Transactions on pattern analysis and machine intelligence, 1986(6): 679—698.

[5] MARTIN D R, FOWLKES C C, MALIK J. Learning to detect natural image boundaries using local brightness, color, and texture cues[J]. IEEE transactions on pattern analysis and machine intelligence, 2004, 26(5): 530—549.

[6] 刘建培, 阎岩, 华祺年, 等. 基于图像处理的圆钢截段长度尺寸测量方法[J]. 内燃机与配件, 2018(14): 239—241.
LIU Jian-pei, YAN Yan, HUA Qi-nian, et al. Measurement method of length and dimension of round steel section based on image processing[J]. Internal combustion engine and accessories, 2018(14): 239—241.

[7] 倪超, 李奇, 夏良正. 粘连棒材图像自动分割计数技术[J]. 数据采集与处理, 2007(1): 72—77.
NI Chao, LI Qi, XIA Liang-zheng. Automatic segmentation and counting technology of adhesion bar image[J]. Data Acquisition and Processing, 2007(1): 72—77.

[8] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition, 2015: 3431—3440.

[9] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImagnNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84—90.

[10] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, 2014: arXiv: 1409.1556.

[11] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition, 2015: 1—9.

[12] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention, 2015: 234—241.

[13] 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.

[14] FLORIAN L C, ADAM S H. Rethinking atrous convolution for semantic image segmentation[C]//Conference on computer vision and pattern recognition(CVPR), 2017: 6—12.

[15] 左纯子, 王征, 张科, 等. 基于改进 DeepLabV3+的煤尘图像分割方法[J]. 工矿自动化, 2022, 48(5): 52—57, 64.
ZUO Jun-zi, WANG Zheng, ZHANG Ke, et al. Coal dust image segmentation method based on improved DeepLabV3+[J]. Industrial and Mining Automation, 2022, 48(5): 52—57, 64.

[16] 汝承印, 张仕海, 张子淼, 等. 基于轻量级 MobileNet-SSD 和 MobileNetV2-DeeplabV3+的绝缘子故障识别方法[J]. 高电压技术, 2022, 48(9): 3670—3679.
YOU Chen-ying, ZHANG Shi-hai, ZHANG Zi-miao, et al. Insulator fault identification method based on lightweight MobileNet-SSD and MobileNetV2-DeeplabV3+[J]. High Voltage Technology, 2022, 48(9): 3670—3679.

[17] 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.

[18] YANG M, YU K, ZHANG C, et al. Denseaspp for semantic segmentation in street scenes[C]//Proceedings of the IEEE conference on computer vision and pattern recognition, 2018: 3684—3692.

[19] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. Computer Science, 2014(4): 357—361.

[20] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition, 2018: 7132—7141.

[21] WANG Q, WU B, ZHU P, et al. Supplementary material for ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 13—19.

责任编辑:陈　芳