

文章编号: 2095-2163(2023)09-0009-04

中图分类号: TP391

文献标志码: A

基于 Mask R-CNN 和 CEDN 的人体实例分割方法

姚 砾, 高 傲, 张乃港, 吴昊然, 万 燕

(东华大学 计算机科学与技术学院, 上海 201600)

摘要: 针对人体实例分割任务中 Mask R-CNN 方法对于边缘分割较差的问题, 本文提出了一种结合轮廓检测算法 CEDN 进行改进的方法。首先, 通过 Mask R-CNN 检测出背景中的人体实例掩码, 通过区域填充获取到精细化人体实例分割结果的方法。首先, 通过 Mask R-CNN 检测出背景中的人体实例掩码; 其次, 利用轮廓检测算法得到精细化人体轮廓, 再通过改进的区域填充算法填充出人体分割掩码, 从而提高人体分割精度。在 LSP 数据集上进行验证, 本算法相较 Mask R-CNN 准确率提高了 4%, 召回率提高了 8%。算法有效的改进了 Mask R-CNN 的分割结果, 改善了人体分割边缘较差的问题, 进一步提升了人体分割的精度。

关键词: 人体实例分割; Mask R-CNN; CEDN

Body instance segmentation method based on Mask R-CNN and CEDN

YAO Li, GAO Ao, ZHANG Naigang, WU Haoran, WAN Yan

(College of Computer Science and Technology, Donghua University, Shanghai 201600, China)

[Abstract] Aiming at the poor edge segmentation of Mask R-CNN method in the human instance segmentation task, a CEDN method was proposed to obtain human contour and then obtain refined human instance segmentation results through region filling. Firstly, Mask R-CNN was used to detect the Mask of the human body in the background, then the refined human contour was obtained by the contour detection algorithm, and then the human body segmentation Mask was filled by the improved region filling algorithm to improve the accuracy of human body segmentation. Compared with Mask R-CNN, the accuracy and recall rate of this algorithm are improved by 4% and 8% respectively. Experiments show that the algorithm can effectively improve Mask R-CNN segmentation results, optimize the problem of the poor edge of human segmentation, and further improve the accuracy of human segmentation.

[Key words] human body instance segmentation; Mask R-CNN; CEDN

0 引言

人体实例分割指对于背景中需要检测出的人体进行提取和分割, 是计算机视觉领域如医学影像、自动驾驶、三维重建等许多任务研究的前提。实例分割是目标检测和语义分割两项任务的结合, 目标检测主要有两种方式: 单阶段方法和两阶段方法。两阶段方法的第一阶段是通过不同的方式产生候选框, 在第二阶段对这些候选框进行分类和微调位置, 获取到检测的结果; 单阶段方法直接对图片进行处理得到检测结果。语义分割则是对输入图像中的每个像素点进行预测, 判断其属于前景还是背景, 从而将前景分割出来。

随着深度学习的发展, 诞生了许多出色的实例分割算法。Long J^[1] 等提出了 FCN (Fully Convolutional

Network) 方法, 将上采样和下采样方法引入图像分割领域中, 针对像素分类有效的对目标进行分割; He 等人^[2] 提出了 Mask R-CNN (Mask Region-based CNN) 算法, 首先提取目标候选框, 然后在候选框内使用 FCN 进行分割, 提升了分割精度, 但 Mask R-CNN 算法对于一些不规则的物体, 例如人、动物等在边缘上分割效果表现的较差; 2018 年, Liu 等^[3] 在 Mask R-CNN 的基础上, 提出了一种两阶段实例分割算法 PANet (Path Aggregation Network), 通过特征金字塔和自适应池化进一步提升了获取特征的能力, 但是在目标边缘上表现的仍然较差。

从以上几种方法能看出来, 由于两阶段的做法先天依赖于第一阶段框选, 导致算法丢失了物体的轮廓信息, 从而很容易在不规则物体的边缘处分割较差。因此本文针对这一问题进行研究, 通过引入

作者简介: 姚 砾 (1967-), 男, 博士, 副教授, 主要研究方向: 图像处理、计算机视觉、深度学习; 高 傲 (1999-), 男, 硕士研究生, 主要研究方向: 图像处理、计算机视觉、深度学习。

通讯作者: 万 燕 Email: winniewan@dhu.edu.cn

收稿日期: 2022-05-30

轮廓信息对分割结果进行优化,进一步提升实例分割的精度。传统的轮廓检测方法通常借助边缘检测算子实现,例如 Laplacian 算子、Canny 算子、Sobel 算子等等,虽然能检测出场景中的边缘信息,但是容易受到噪声干扰,检测出冗余的边缘信息。Yang J 等人^[4]提出轮廓检测方法 CEDN (Convolutional Encoder-Decoder Network) 不仅可以有效的提取出目标轮廓,而且受到噪声干扰较小,因此本文基于 CEDN 算法进行改进,用于提取人体实例轮廓,优化 Mask R-CNN 的分割结果。

1 基于 Mask R-CNN 和 CEDN 的人体实例分割方法

1.1 总体方法介绍

虽然 Mask R-CNN 算法在人体实例边缘处分割结果较为粗糙,但是对整体的分割仍然是有效的,因此可以利用 Mask R-CNN 对人体进行初步分割,得到大致的形状;利用 CEDN 算法对图像中的轮廓信息进行检测;通过 Mask R-CNN 分割的掩码提取出人体轮廓;最后,对人体轮廓进行区域填充,便可以获取到较为精准的人体实例分割结果。本文方法流程图如图 1 所示。



图 1 本文实例分割方法流程图

Fig. 1 The flowchart of the instance segmentation method in this article

本文通过像素级与运算将人体掩码和轮廓掩码融合,为了防止人体轮廓出现断裂,需要将人体掩码膨胀,完整覆盖轮廓掩码后再进行与运算。研究发现,膨胀次数介于 15 次和 20 次之间效果最好,可以得到较为完整的人体轮廓。膨胀操作如图 2 所示。

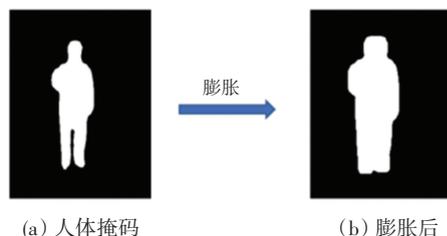


图 2 膨胀操作

Fig. 2 Dilated progress

简单的像素级与运算融合,人体轮廓边缘难免会出现毛刺,如图 3 所示。

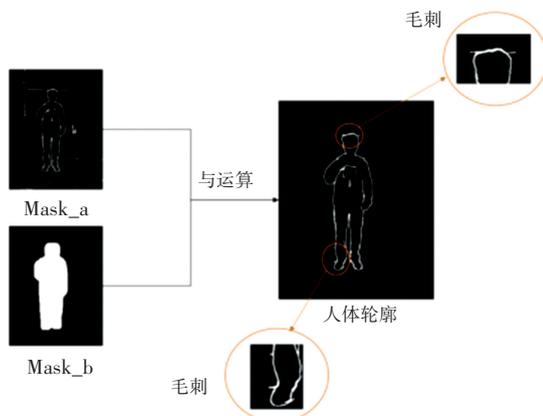


图 3 人体轮廓毛刺

Fig. 3 Body contour burr

后续区域填充需要光滑的轮廓,需要通过形态学手段进行光滑处理,去除掉毛刺。本文观察人体轮廓发现,毛刺宽度超过 2 个像素,人体轮廓如图 4 所示。

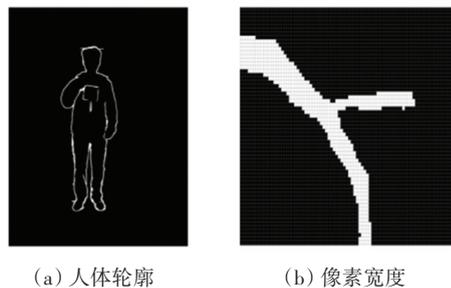


图 4 轮廓宽度

Fig. 4 Contour width

而超过 2 个像素宽度的毛刺使用一般的形态学开闭运算难以细化和消除。因此本文设计了一种基于形态学的毛刺去除方法,有效的去除掉毛刺信息。首先,使用骨架提取方法,将轮廓和毛刺细化成像素宽度为 1 的骨架;由于毛刺多是水平方向,因此通过线扫描法水平线性扫描获取到水平形状人体掩码,利用形态学开运算有效的去除毛刺;通过和骨架的融合,便可以得到去除毛刺后的人体轮廓。毛刺去除过程如图 5 所示。

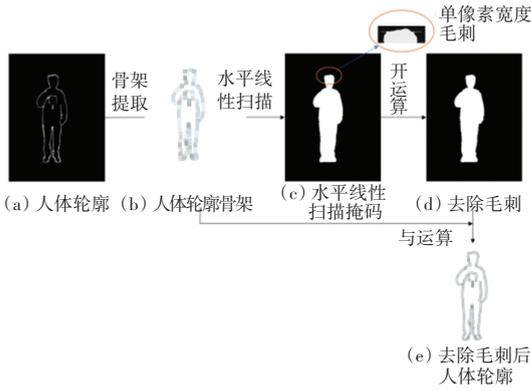


图 5 去除毛刺过程
Fig. 5 Deblurring

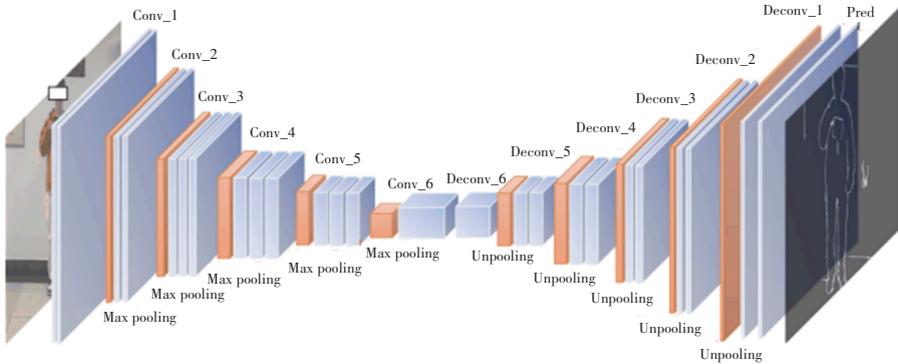


图 6 轮廓检测算法网络结构图

Fig. 6 The architecture of the CEDN algorithm network

本文将改进后的轮廓检测方法和 Canny 算子, Laplacian 算子以及 Sobel 算子的检测效果进行了对比,结果如图 7 所示,可见本文改进后的轮廓检测方法去除了大量的服装纹理等信息带来的冗余轮廓。

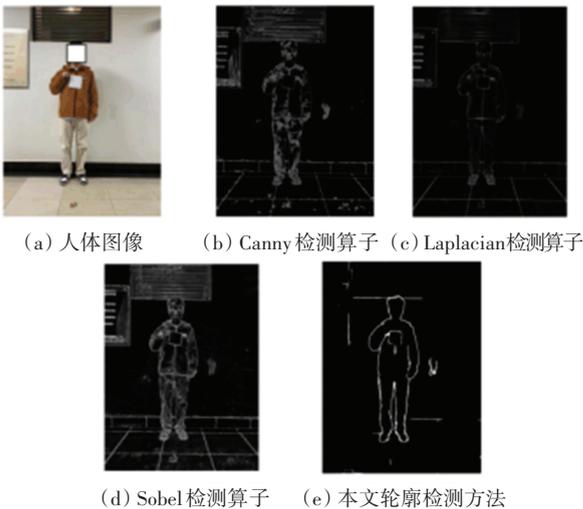


图 7 边缘检测技术比较

Fig. 7 Comparison of the edge detection technology

1.3 基于扫描线法的区域填充

本文区域填充方法的流程图如图 8 所示。获取到人体实例轮廓后,需要对轮廓进行填充以获得完整的人体实例分割图像。由于轮廓内部服装纹理信息的干

扰,普通的区域填充算法,例如连通域算法、泛洪填充法等等填充效果较差。本文提出一种基于扫描线法的区域填充算法,对于一个待填充人体轮廓图,从左到右、从上到下以及其对应的逆方向分别扫描,当遇到像素值等于 255 时,即认为到达了轮廓边界,本次扫描停止,转向从下一行或者下一列开始重复扫描。

1.2 基于 CEDN 的轮廓检测

CEDN 算法虽然能够有效的提取出目标的轮廓,但鲁棒性较差,检测出的人体轮廓容易出现断裂的问题。因此本文基于 CEDN 算法的网络结构做了一定的调整,增加解码器的卷积层个数,以增加解码器解码能力,有助于避免轮廓出现断裂的现象。轮廓检测算法网络结构如图 6 所示。

扰,普通的区域填充算法,例如连通域算法、泛洪填充法等等填充效果较差。本文提出一种基于扫描线法的区域填充算法,对于一个待填充人体轮廓图,从左到右、从上到下以及其对应的逆方向分别扫描,当遇到像素值等于 255 时,即认为到达了轮廓边界,本次扫描停止,转向从下一行或者下一列开始重复扫描。

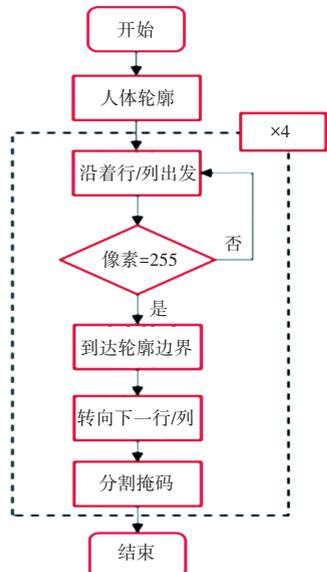


图 8 区域填充流程图

Fig. 8 Region filling flow chart

由于内部纹理以及边缘毛刺的原因,传统的区域填充法效果并不理想,本文采用结合形态学的扫描线法区域填充方法,有效的对人体轮廓进行填充,分割出人体实例。

将改进后的区域填充方法和传统的泛洪填充法和线扫描法进行了对比,结果如图9所示。可见本文改进的区域填充方法有效的填充了轮廓内部区域。

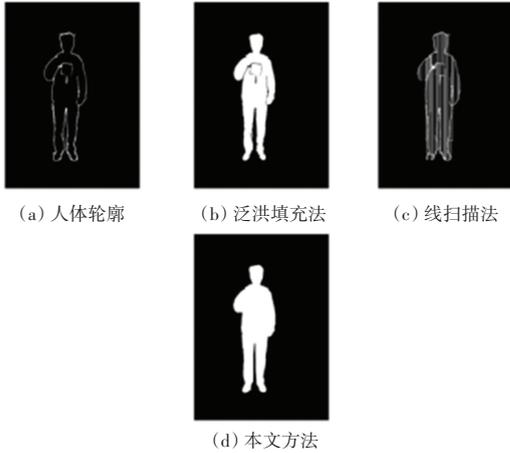


图9 区域填充算法比较

Fig. 9 Comparison of the region filling

2 实验结果与分析

2.1 数据集和评价指标

本文在 LSP (Leeds Sports Pose) 数据集上进行实验和测试,并分别计算准确率 (*Precision*) 和召回率 (*Recall*) 评价本文算法的性能,准确率和召回率数学定义分别为式(1)和式(2):

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

其中, *TP* 表示被正确预测为正样本的数量; *FP* 表示负样本被预测为正样本的数量; *FN* 表示正样本被预测为负样本的数量。

2.2 定量分析

以 Mask R-CNN 算法为比较对象,评价指标见表1。可见与 Mask R-CNN 算法相比,本文方法准确率提高了4%,召回率提高了8%。

表1 评价指标

Tab. 1 Evaluation metrics

评价指标	Mask R-CNN	本文方法
精确率 <i>P</i>	0.87	0.91
召回率 <i>R</i>	0.80	0.88

2.3 定性分析

通过在 LSP 数据集上选择图片进行分割测试,结果如图10所示。本文方法在人体实例的头部、手肘、脚部等地方相比较 Mask R-CNN 的分割结果更加精细。实验结果验证了本文方法的有效性,即通过轮廓检测算法以及形态学、区域填充手段的引入使得本文算法不仅在分割精度相对于 Mask R-CNN 得到了提升,而且进一步优化了目标边缘。

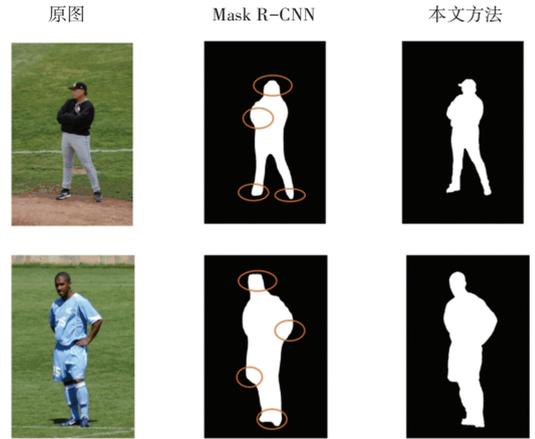


图10 Mask R-CNN方法与本文分割结果对比

Fig. 10 Segmentation comparison of Mask R-CNN method and ours

3 结束语

本文针对 Mask R-CNN 算法在人体实例分割中存在分割边缘较差、分割精准度较低等问题,轮廓检测算法 CEDN 提取出人体轮廓,然后通过改进的区域填充算法填充人体轮廓,从而得到人体分割掩码。本文算法有效的利用轮廓信息优化了人体边缘,改善了 Mask R-CNN 算法分割出的人体边缘粗糙的问题。在 LSP 数据集上实验,本文方法较 Mask R-CNN 算法精确率提高了4%,召回率提高了8%。

参考文献

- [1] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640-651.
- [2] HE K, GKIOXARI G, DOLLÁR P, et al. Mask r-cnn [C]// Proceedings of the IEEE international conference on computer vision. 2017: 2961-2969.
- [3] 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.
- [4] YANG J, PRICE B, COHEN S, et al. Object contour detection with a fully convolutional encoder-decoder network [C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 193-202.