

基于智能视觉的机械零件图像分割技术

洪庆¹, 宋乔¹, 杨晨涛¹, 张培¹, 常连立²

(1. 南京理工大学 机械工程学院, 江苏 南京 210094; 2. 北京航天新风机械设备有限责任公司, 北京 100854)

摘要:为有效获取零件特征以提高现代生产智能化、精密化水平,基于智能视觉的机械零件分割研究起着关键作用。针对航天机器人等装配车间流水线零部件智能感知问题,研究基于智能视觉的零部件图像分割算法,实现机械零部件分割识别。基于 Deeplabv3 图像分割算法提出一种增加自定义 Encoder-Decoder 特征提取模块的网络结构 Deeplabv3-d,采用掩膜标记特征区域,基于该改进网络结构,采用 mobileNet 和 Resnet101 两种骨干网络进行零件图像分割对比实验,证明了该图像分割算法在零件图像分割应用领域的实用性。

关键词:智能制造;深度学习;图像分割;Encoder-Decoder;Deeplabv3

中图分类号:TP391.41 **文献标志码:**A **文章编号:**1671-5276(2020)05-0203-04

Image Segmentation Technology of Mechanical Parts Based on Intelligent Vision

HONG Qing¹, SONG Qiao¹, YANG Chentao¹, ZHANG Pei¹, CHANG Lianli²

(1. School of mechanical engineering, Nanjing University of Science and Technology, Nanjing 210094, China;

2. Beijing Aerospace Xinfeng Mechanical Equipment Co., Ltd., Beijing 100854, China)

Abstract: To obtain the features of parts effectively and improve the intelligence and precision of modern production, this paper researches on mechanical parts segmentation based on intelligent vision plays a key role. Aiming at the problems existing in intelligent perception of assembly line parts in aerospace robot and other assembly workshops, the image segmentation algorithm based on intelligent vision is studied, which is used to realize the segmentation and recognition of the mechanical parts. Based on the Deeplabv3 image segmentation algorithm, a network structure, Deeplabv3-d is proposed in which a user-defined Encoder-Decoder feature extraction module is added. The mask is used to mark the feature area. Based on the improved network structure, two kinds of backbone networks: mobileNet and Resnet-101 are provided, it is proved that the algorithm is of practicality in the field of image segmentation.

Keywords: intelligent manufacturing; deep learning; image segmentation; Encoder-Decoder; Deeplabv3

0 引言

随着计算机技术的飞速发展,图像分析处理已成为视觉的重要研究方向,在制造装配领域有着广阔的应用前景。

传统视觉算法对于特征十分敏感,而实际的工作环境复杂,光线条件恶劣,传统算法无法满足实际生产需要。基于深度学习的图像处理算法相对于传统算法具有更高的准确性和稳定性,在制造业应用更加广泛。苏朝阳等基于机器视觉,利用图像处理算法组件分别配置图像处理链,实现了对电容器外观底部露白缺陷和引脚变形缺陷的检测^[1]。周振等针对图像目标受环境影响难识别目标的问题,基于深度图像设计并实现了一种轮廓提取及特征识别的方法^[2]。魏中雨等针对制造中零件识别定位以及装配检测效率低、准确度差等问题,提出了一种基于深度神经网络的零件装配检测方法^[3]。张辉等提出一种基于机

械臂和机器视觉的表面质量检测方法^[4]。

在现有技术条件下,预设特征、姿势再进行工程学分析的方法越来越难以满足现代制造的装配质量及效率要求。如何实现装配实时工程学分析已成为一个迫切需要解决的问题^[5]。为提高实时装配分析的可行性,本文将针对其中零件图像分割技术展开研究。

1 技术路线

本文实现机械零件图像分割的技术路线如下:

1) 根据制造过程中零件准确率及速度要求,选取 Deeplabv3^[6-7] 算法模型作为本文机械零件图像分割的基础算法。

2) 机械零件数据集的准备。收集大量机械零件相关图片,并利用标注工具生成图像掩膜标签,将数据分为训练集、验证集、测试集。

3) 基于 Deeplabv3 算法模型,加入自定义 Encoder-

基金项目:国防基础科研重点资助项目(JCKY2017204B053);国防预先研究资助项目(41423010203);中央高校自主科研基金资助项目(30919011208)

第一作者简介:洪庆(1997—),男,安徽芜湖人,硕士研究生,研究方向为智能制造视觉系统开发。

Decoder 特征提取模块,构建改进后的 Deeplabv3 算法,并命名为 Deeplabv3-d。

4) 基于自建机械零件数据集,完成 Deeplabv3-d 的训练与测试,并根据测试结果评估算法的实用性。

5) 基于 Deeplabv3-d 算法,更换不同骨干网络 (mobileNet、Resnet-101),在自建机械零件数据集上进行性能对比测试。

2 Deeplabv3-d 算法原理

2.1 空洞卷积(膨胀卷积)

图像分割过程中,初始图像先用卷积网络层进行特征卷积提取,生成卷积后的特征图,再由池化层进行最大或平均特征的提取,之后图像尺寸变为原尺寸的一半,再经过反卷积网络层将尺寸恢复至原始尺寸,保证提取特征、扩大感受野的同时,还原得到的预测像素信息。

因此图像分割模块主要包括两个部分:1) 通过卷积池化层(即下采样层)提取图像特征并保证较大的感受野信息;2) 通过反卷积层(即上采样层)扩大图像尺寸并还原需要得到的预测信息。在下采样—上采样过程中,很容易出现数据信息丢失、空间结构紊乱、层级结构信息丢失等现象,例如小目标检测时无法重建完整信息。受制于该问题,图像分割算法长期处在瓶颈期,而空洞卷积^[8]的出现使该问题得到了较好的解决。

图 1 是常规卷积和空洞卷积的卷积操作图^[8],通过对比可知,相比于常规卷积的一次常规感受野卷积,空洞卷积对更大的感受野进行卷积操作,使卷积后的每一个特征点都涵盖了更广的信息。

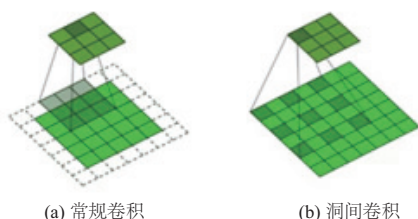


图 1 卷积操作图

图 1(b) 为空洞=2 时的空洞卷积操作图,一般性的卷积映射公式见式(1):

$$y[i] = \sum_k x[i + rk] w[k] \quad (1)$$

其中: r 表示卷积步幅即空洞,标准卷积是比率 $r=1$ 的一种特殊情况。通过更改步幅值(空洞)可自适应地修改卷积层提取特征的感受野。

2.2 深度可分离卷积

深度可分离卷积^[9]思想是将标准卷积层分解为深度卷积层、逐点卷积层(即 1×1 卷积),在不增加网络计算复杂度的基础上,大大增加了网络深度以及网络泛化、拟合能力。具体而言,深度卷积针对每个输入通道独立执行空间卷积,而点向卷积用于组合深度卷积的输出^[10]。这是一种轻量级的卷积方式,将传统的多通道卷积模型分解为

逐通道卷积+逐点卷积,相比起常规的卷积,其运算成本和计算量都大为降低,实现算法复杂度优化,大大加快了训练速度。

2.3 Deeplabv3-d

Encoder-Decoder^[10] 特征提取模块,在深度神经网络中常用于图像分割任务。普通卷积网络层能够基于卷积、池化方式编码多维度信息,通过使用不同尺度、步长的卷积核实现多种速率或多个感受野的变换,并利用最大池化操作来探查传入特征^[7],而 Encoder-Decoder 模块能够通过特征空间的反向传播来获取更为准确的对象边界。

为优化 Deeplabv3 图像分割性能,在原本 v3 模型中添加 Encoder-Decoder 特征提取模块,减少其数据信息丢失、空间结构紊乱、层级结构信息丢失的可能性,改善其对象边界分割效果。再将深度可分离卷积理论应用于 Encoder-Decoder 模块,从而在不增加计算复杂度的基础上构建出更深网络的 Encoder-Decoder 模块,使得 Deeplabv3 算法网络泛化拟合能力大大增加,本文将改进后的 Deeplabv3 命名为 Deeplabv3-d。

改进的 Deeplabv3,使用了空间金字塔池化和 Encoder-Decoder 结构。

本文提出的模型 Deeplabv3-d 包含来自编码器模块的丰富语义信息,而对象边界提取由简单而有效的 Decoder 恢复模块进行,Encoder 模块允许通过应用空洞卷积以任意分辨率提取卷积特征。在最后特征图中的编码丰富了语义信息,而对于网络中的卷积池化操作引起的对象边界小部分信息丢失问题,可以通过应用空洞卷积提取更密集的特征图来减弱。

因考虑到有限的 GPU 内存,提取出比输入小 8 倍甚至 4 倍的输出特征图在计算上并不可取。以 Resnet101^[11] 为例,当利用空洞卷积提取比输入分辨率小 16 倍的输出特征时,必须对其特征进行扩展,而要求输出特征比输入小 8 倍,则卷积块会增加大量的网络参数。

Encode-Decoder 模块有助于在 Encoder 通道中进行更快的计算,并逐渐在 Decoder 通道中恢复清晰的对象边界。为结合两种方法各自的优点,本文通过合并多维度上下网络层特征信息来丰富 Encoder-Decoder 网络模块,具体结构如图 2 所示。

图 2 即为完整的 Deeplabv3-d 结构。Deeplabv3 加入自定义 Encoder-Decoder 模块后,将采用 mobileNet 与 Resnet101 进行拟合效果比对。

3 Deeplabv3-d 训练及分析

为使模型系统更快拟合而自建机械零件数据集,使用批量梯度下降法,将数据集分为 100 个批次,设置周期为 50,迭代次数为 600,进行训练迭代。

训练过程损失值变化如图 3 所示,可发现模型训练损失在第 200 次迭代后逐渐趋于平稳,最终达到 0.014 53。根据实际曲线波动情况得知,模型在整个训练过程中仅出现少量震荡,拟合效果较好。

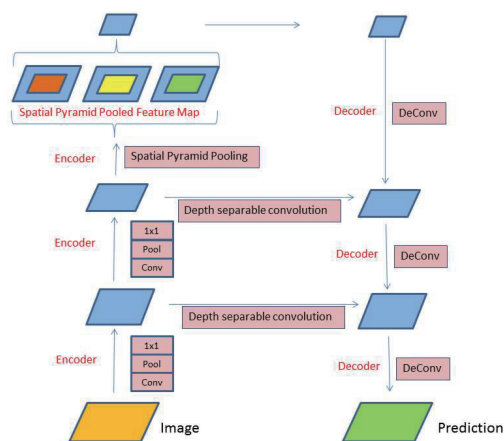


图2 自定义 Encoder-Decoder 模块

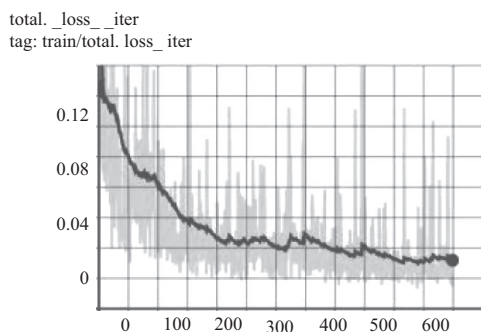


图3 训练 loss-迭代次数图像

4 Deeplabv3-d 零件图像分割测试

观察图4所示的单个零件分割效果、图5所示的多个零件堆叠分割效果可知,Deeplabv3-d 对机械零件图像的分割效果较好,与零件轮廓拟合度较高。

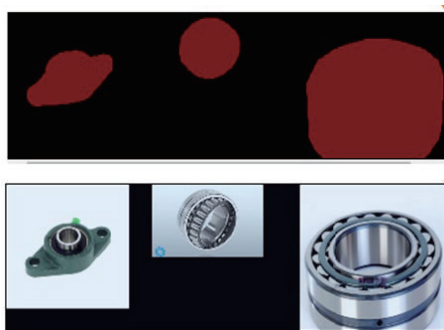


图4 单个零件分割效果图

综合分析图4和图5可得,不同环境下,Deeplabv3-d算法网络对单个零件或多个零件堆叠都具有良好、直观的分割效果,应用于实际工业场景中具有一定的可行性。同时,Deeplabv3-d在测试集上的预测准确度 Acc 和平均交并比 m_{iou} 随训练周期变化曲线如图6、图7所示。



图5 多个零件堆叠分割效果图

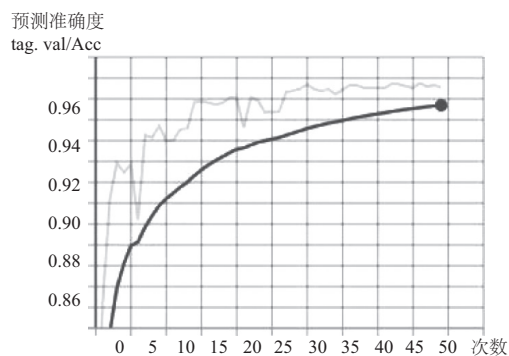
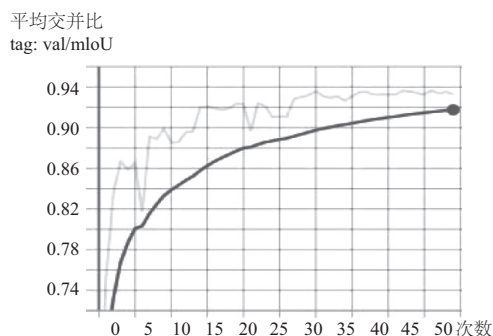


图6 测试集 Acc 随训练周期变化曲线

图7 m_{iou} 随训练周期变化曲线

分析图6、图7中 Acc/ m_{iou} 随周期的变化曲线,可得出模型的预测准确度和平均交并比经过震荡上升,在第30个训练周期后逐渐趋于平滑,最终达到90%以上,在分割精度上有较好的表现。

为对比 Deeplabv3-d 不同主干网络中的模型拟合效果,采用两种骨干网络进行验证。在表1中记录了在 mobileNet 和 Resnet101 两种骨干网络下的图像分割效果。两种骨干网络平均 m_{iou} 值都超过90%,均展现出较好的分割性能。

表1 不同骨干网络下的 Deeplabv3+评估

骨干网络	$m_{iou}/\%$
Deeplabv3+mobileNet	93.57
Deeplabv3+Resnet101	94.34

5 结语

本文针对机械零件图像分割,通过建立机械零件数据集,并向 Deeplabv3 算法网络中加入集成空间金字塔、空洞卷积、深度可分离卷积结构的 Encode-Decoder 特征提取模块,完成了改进网络 Deeplabv3-d 的构建。在实际单零件或多零件堆叠图像的分割效果测试中,Deeplabv3-d 取得了较好的分割效果,证明了在实际环境中零件图像分割算法的可行性,对实际工业生产具有较大意义。

参考文献:

- [1] 苏朝阳,马万太,黄磊,等. 基于图像处理库的机器视觉检测系统的研究[J]. 机械制造与自动化,2019,48(1):193-196.
- [2] 周振,杜姗姗. 基于 kinect 深度图像的目标定位与识别[J]. 机械制造与自动化,2016,45(4):173-176.
- [3] 魏中雨,黄海松,姚立国. 基于机器视觉和深度神经网络的零件装配检测[J]. 组合机床与自动化加工技术,2020(3):74-77,82.
- [4] 张辉,余厚云,李克斌. 基于机械臂的涡轮壳零件表面质量视觉检测[J]. 计算机与现代化,2020(7):121-126.
- [5] 陈晓春,林博溢,孙乾,等. 基于运动区域差分与卷积神经网络的动作识别[J]. 计算机工程,2019,45(12):274-280,293.

- [6] 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.
- [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 & Machine Intelligence, 2016(99):834-848.
- [8] 曲长波,姜思瑶,吴德阳. 空洞卷积的多尺度语义分割网络[J]. 计算机工程与应用,2019,55(24):91-95.
- [9] 马立,巩笑天,欧阳航空. Tiny YOLOV3 目标检测改进[J]. 光学精密工程,2020,28(4):988-995.
- [10] ANDRIJASA M F. Deep learning with encoder-decoder architecture for exchange currency rates model predictions[J]. Journal of Physics: Conference Series, 2019,1402(6):066098-066102.
- [11] YANG Lu, SONG Qing, WU Yingqi, et al. Attention inspiring receptive-fields network for learning invariant representations [J]. IEEE Transactions on Neural Networks and Learning Systems, 2019,30(6):1744-1755.

收稿日期:2020-07-29

(上接第 202 页)

用这些点对求解转换矩阵并完成匹配,平均偏差为 0.029 mm,足以满足装配周期较长的大尺寸装配摄影测量变形检测的需求。同时,本文的算法也能应用于其他需要标志点匹配的情形,例如基于标志点匹配的点点云拼接。为保证匹配结果的准确性,本文对未发生形变的点对筛选过于严格,被筛选出的点在邻域内几乎不发生形变,今后将进一步改进,以适应更多稀疏标志点匹配的应用场景。

参考文献:

- [1] 隆昌宇,郝继贵,郭寅,等. 基于非参数测量模型的摄影测量方法研究[J]. 光学学报,2014,34(12):201-209.
- [2] 蒋悦,黄宏光,舒勤,等. 高维正交子空间映射的尺度点点云配准算法[J]. 光学学报,2019,39(3):290-300.
- [3] ZHANG S, WANG H, GAO J G, et al. Frequency domain point cloud registration based on the fourier transform[J]. Journal of Visual Communication and Representation, 2019(5):170-177.
- [4] HALBER M, FUNKHOUSER T. Computer Vision and Pattern Recognition (CVPR)[M].[S.I.]: IEEE, 2017.
- [5] CHEN C S, HUNG Y P, CHENG J B. RANSAC-based darces: a new approach to fast automatic registration of partially overlapping range images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999,21(11):1229-1234.
- [6] 熊风光. 三维点点云配准技术研究[D]. 太原:中北大学,2018.
- [7] RUSU R B, BRADSKI G, THIBAU R, et al. Fast 3d recognition and pose using the viewpoint feature histogram [C]//2010

IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2010:2155-2162.

- [8] RUSU R B, BLODOW N, BEETZ M. Fast point feature histograms (FPFH) for 3D registration[C]//2009 IEEE International Conference on Robotics and Automation. IEEE, 2009:3212-3217.
- [9] 刘剑,白迪. 基于特征匹配的三维点点云配准算法[J]. 光学学报,2018,38(12):240-247.
- [10] 黄源,达飞鹏,陶海路. 一种基于特征提取的点点云自动配准算法[J]. 中国激光,2015,42(3):250-256.
- [11] 翁金平,程筱胜,崔海华,等. 面向三维测量的空间标志点高稳定实时匹配方法[J]. 激光与光电子学进展,2016,53(2):103-110.
- [12] RUSU R B, MARTON Z C, BLODOW N, et al. Persistent point feature histograms for 3D point clouds [C]//Baden-Baden, Germany: Proc 10th Int Conf Intel Autonomous Syst, 2008:119-128.
- [13] MIKOLAJCZYK K, SCHMID C. A performance evaluation of local descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005,27(10):1615-1630.
- [14] WEILIANG L U, JIANG K, LIN J. Multi-view 3D data registration based on global control codeless points[J]. Opto-Electronic Engineering, 2014(5):57-62.

收稿日期:2019-07-04