

# 机械零件识别方法

肖奎<sup>1,2</sup>

(1. 陕西理工大学机械工程学院, 陕西汉中 723000;  
2. 陕西省工业自动化重点实验室, 陕西汉中 723000)

摘要: 针对装配场景中零件之间相互遮挡和不同位姿的影响导致零件识别效果不佳的问题, 提出一种基于改进的 Mask R-CNN 的复杂场景零部件识别方法。首先, 采用 ResNet101 网络增大网络层数来加深网络增强网络特征提取能力; 其次, 改进融合特征的特征金字塔网络结构, 提升高低层网络之间传递效率, 增强对于特征的捕获能力。通过在自制数据集的实验, 发现与原始 Mask R-CNN 相比, 该方法检测精度有明显提升, 为复杂环境的零件识别提供解决方案。

关键词: 识别与定位; 目标检测; 深度学习; 实例分割

零件作为机械生产的重要组成部分, 其种类复杂, 数量繁多, 对零件的误识别会导致生产效率降低和资源浪费, 甚至会导致事故。因此, 机械零件的准确识别是当前研究的热点问题。

传统人工识别效率低下, 受主观影响较大; 利用相机获取零件图像, 并利用机器学习算法对目标进行感知识别与定位可以克服这些问题, 因此, 零件识别与定位成为机器人视觉任务研究中的重要组成部分。传统方法对特征算子依赖性强, 对检测条件和目标的状态要求严苛, 针对不同任务要进行新的设计。

随着计算机性能的不不断提升和深度学习技术的发展, 神经网络在零件识别领域取得了显著优势。本文改进的 Mask R-CNN 实例分割模型, 通过改进网络的主干网络和特征融合网络的结构, 在自制数据集上进行实验发现, 本文网络可以明显提高检测精度, 基本可以实现复杂环境下零件的准确识别。

## 一、数据集与方法

### (一) 数据集制作

实验采用自己制作的零件图片数据集, 样本涵盖螺栓, 螺母两种主要机械零件, 主要包含零件堆叠的情况下的图片数据集。通过数据增强的技术手段, 采用滤波, 加噪声等方法模拟生产过程中复杂的工作环境。不仅能够大幅扩大数据集规模, 还能保留图像特征, 增加样本训练的多样性, 提高网络的泛化能力。将数据集增强后, 与原始图像融合, 使模型的鲁棒性更加稳定。图像分辨率 640×640。使用 Labelme 标注软件将数据集打标获得 JSON 格式文件。



图1 数据集样本

### (二) 主干网络选择

Mask R-CNN 的主干网络一般为 ResNet50 和 ResNet101, 区别主要在于网络的深度, 网络深度决定了网络的表达能力, 网络越深学习能力越强。为了增强网络的特征提取能力, 本文采用 ResNet101 主干网络。

ResNet101 通过引入残差单元和在网络中增加直连通道, 减少了参数数量, 增加了网络对特征的学习能力, 具有参数少和识别效果优秀等特点, ResNet101 总共由 101 层组成, 其中包括第一层 7×7 的卷积层, 4 个阶段, 最后一层的全局平均池化和一层全连接层。其中的每个阶段都含有几个残差块。每个残差块都包含两个 3×3 的卷积层, 后面有一个批量归一化层和激活函数层, 并且残差块之间也有批量归一化层和激活函数。ResNet101 允许网络在学习过程中直接关注输入与输出的残差, 增强了网络的特征提取能力, 提高网络的识别效率。

### (三) 特征金字塔网络

特征金字塔网络 (FPN) 融合特征之前, 主干网络的不同阶

段先通过卷积进行目标的特征提取，然后特征金字塔网络对上层的特征图进行上采样，对下面一层特征图进行下采样从而实现特征融合，融合后的特征图再经过相同操作与下层特征融合。因为不同阶段感受野大小不同，包含的语义信息也不同，高层的卷积层特征语义信息多，低层的卷积层位置信息多，所以，对他们进行直接特征相加操作，会减弱网络的特征提取能力，影响网络的检测效果。为了更好的在不同层上面输出相应的目标，提高网络的特征融合的效率加速网络的运行，提升网络的检测性能，本文提出一种优化的特征金字塔网络。

为了更好的在不同层上面输出相应的目标，提高网络的特征融合的效率加速网络的运行，提升网络的检测性能，本文提出一种优化的特征金字塔网络。改进的网络通过添加横向连接，自上而下和自下而上的路径，每个特征图上使用  $1 \times 1$  卷积与相邻路径进行特征融合，在每个特征图上通过卷积投影特征并与相邻路径进行特征融合。增加金字塔通道作为自下而上的信息流，通过采用最接近邻插值方式对高层特征图进行上采样然后通过卷积与低层特征融合。和传统特征金字塔网络相比，改进后的网络能将更多的语义信息融合，提高小零件的识别精度，减小零件堆叠造成的语义信息缺失对零件识别的影响，从而减小了零件的误识别

概率。

(四) 评价指标

为评价模型性能，实验采用检测精度 Precision，召回率 Recall 和平均检测精度 mAP 指标来衡量。其中，

检测精度分为 seg\_mAP 和 Bbox\_mAP，分别表示网络的检测和分割精度。mAP 值越高，代表模型检测性能越好。公式为：

$$\text{precision} = \frac{TP}{TP+FP} \tag{1}$$

$$\text{recall} = \frac{TP}{TP+FN} \tag{2}$$

$$F1 = 2 \times \frac{\text{precision} \cdot \text{recall}}{(\text{precision} + \text{recall})} \tag{3}$$

$$AP = \int_0^1 p(r) dr \tag{4}$$

$$mAP = \sum_{c=1}^C \frac{AP}{C} \tag{5}$$

式中：TP 为正确识别并标定目标的数目；FP 为错误识别但标定目标的数目；FN 为错误识别且未能标定目标的数目。

二、结果与分析

(一) 试验对比

为了提高模型的检测精度，本文设计改进特征融合网络和主干网络，提升了网络对复杂环境目标的检测精度。本文对各个模型的参数进行对比，评估结果如表 1。

表 1 模型实验

通过排	Recall/%	precision/%	F1/%	Seg_mAP/%	Bbox_mAP/%	模型大小 /M	FPS/Gflops
Mask R-CNN	89.2	86.6	87.9	87.8	96.7	43.75	258.17
Mask R-CNN+r101	92.5	89.2	90.8	89.2	97.8	62.74	334.24
Mask R-CNN+GN	90.4	88.4	89.3	88.4	97.7	45.79	465.04
Mask R-CNN+M-FPN	91.5	89.0	90.2	88.9	97.6	47.05	356.32
Ours	94.2	91.3	92.7	92.1	98.0	66.03	480.34

本文算法相对于其他的算法，在大小差不多的情况下，精度和召回率都有明显的提高。所以，本文零件识别算法具有一定效果。

三、结论

零件识别为零件装配中的关键技术，本文采用 Mask R-CNN 实例分割模型为基础，通过改进主干网络和特征融合网络的结构来增强网络对特征信息的提取和利用，从而实现零部件的高精度检测与分割。但本文识别方法仍有不足，下一步需要针对零件识别进行专门训练；是对网络进行轻量化改进，提高识别速度。

参考文献：

[1] 王耀南，江一鸣，姜娇等. 机器人感知与控制关键技术及其智能制造应用 [J]. 自动化学报，2023，49（03）：494-513.  
 [2] 田中可，陈成军，李东年等. 基于深度图像的零件识别及装配监测 [J]. 计算机集成制造系统，2020，26.  
 [3] 孙书魁，范菁，孙中强等. 基于深度学习的图像数据增强研究综述 [J]. 计算机科学，2023.