

DOI:10.19651/j.cnki.emt.2108409

基于 Mask R-CNN 实例分割的 机械零件识别方法研究*

臧春华¹ 周介祺² 刘桂雄²

(1.广东省珠海市质量计量监督检测所 珠海 519060; 2.华南理工大学 机械与汽车工程学院 广州 510640)

摘要: 零件识别是机械部件装配、装箱的重要基础,人工识别效率低,传统机器视觉检测要求高、场景单一。提出一种基于深度学习机器视觉的机械零件识别方法,通过加入 PointRend 模块提升原始 Mask R-CNN 实例分割模型的检测精度;针对相似度高零件设计类别细分方法,通过尺寸估算与特征匹配,较好地解决由于数据增强图像缩放造成的尺寸特征丢失问题。采集 25 种不同零件进行识别实验,结果表明,该方法可有效提升机械零件的识别准确率,算法对相似零件识别准确率达 100%,较原始 Mask R-CNN 方法提升 11.51%。并且该方法可推广到其他具有相似特征目标的识别任务中。

关键词: 目标识别;实例分割;深度学习;特征匹配

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 460

Research on recognition method of mechanical parts based on Mask R-CNN instance segmentation

Zang Chunhua¹ Zhou Jieqi² Liu Guixiong²

(1. Guangdong Zhuhai Supervision Testing Institute of Quality and Metrology, Zhuhai 519060, China;

2. School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: Part recognition is an important basis for the assembly and packing of mechanical components. The efficiency of manual recognition is low. The traditional machine vision inspection requires high and the scene is single. This paper proposes a machine part recognition method based on deep learning machine vision. The detection accuracy of the original Mask R-CNN instance segmentation model is improved by adding the PointRend module. The category subdivision method is designed for parts with high similarity, through size estimation and feature matching, it can better solve the problem of size feature loss caused by data-enhanced image scaling. Collecting 25 different parts for recognition experiments, the results show that the method can effectively improve the recognition accuracy of mechanical parts, and the algorithm can recognize similar parts with an accuracy of 100%, which is 11.51% higher than the original Mask R-CNN method. And the method can be extended to other recognition tasks with similar features.

Keywords: target recognition; instance segmentation; deep learning; feature matching

0 引言

机械生产、装配、装箱过程中使用零件种类繁多、数量巨大,零件识别错误会延误工期、增加浪费,甚至造成安全事故。现有的人工识别方法效率较低,基于机器视觉识别方法能够实现无接触、无间断、可溯源的快速检测,能够有效提高生产效率。现有机器视觉识别方法主要分为模板匹

配、深度学习两种^[1]。

基于模板匹配识别方法重点在于图像特征描述方法设计,通过获取目标特征向量并与库中模板进行比较,使用欧氏距离、支持向量机(support vector machine, SVM)等方法实现识别、分类,常用特征描述方法有角点检测、Hu 不变矩、边缘检测等^[2]。文献[3]通过拍摄两个角度图像,在粗粒度、细粒度特征层次上提取形状因子、不变矩,并通过

收稿日期:2021-10-20

* 基金项目:广东省市场监督管理监督局项目(2021CZ17)、广东省重点领域研发计划项目(2019B010154003)资助

人机交互引导与样本库中图像比对,实现高度相似航空钣金件准确区分;文献[4]提出融合单应性约束尺度不变特征变换(scale-invariant feature transform, SIFT)特征匹配轴承滚子检测方法,对图像局部进行 SIFT 特征提取,并使用快速近似最邻近(fast approximate nearest neighbor, FANN)算法进行特征匹配、筛选;文献[5]设计一种 SIFT-SVM 方法识别轴承盖,使用 SIFT 提取特征矩阵,得到词袋模型(bag of words, BoW)描述,再使用 SVM 进行分类。模板匹配高度依赖人为设计特征算子,对检测条件、目标状态要求较高,场景适应性差,针对不同检测任务需要重新设计。深度学习识别方法通过多层卷积神经网络进行目标特征自动提取,同时实现目标区域提示与类别划分,近年来被广泛应用于图像识别领域,常用模型有基于候选区域的卷积神经网络(regions with convolutional neural network features, R-CNN)系列、一阶段目标检测网络(you only look once, YOLO)系列^[6]。文献[7]使用多分支特征融合卷积神经网络(multi-branch feature fusion convolutional CNN, MFF-CNN)进行主轴盖识别分类,识别率达 91.6%,并能适应多种光照条件;文献[8]在抓取任务中使用级联式 Faster R-CNN 识别目标物体,并设计二分旋转方法,指导机械手进行目标抓取;文献[9]通过对 YOLO v4 模型改进,在边界框识别基础上使用边缘检测算法,获得目标最小外接矩形,有效提高识别准确率和速度。可以看出,深度学习通过多层卷积保留更多特征信息,能归纳出更一般规律,适应更多、更复杂检测任务。目前主要深度学习识别方法对于特征差异大的零件识别效果良好,但由于数据增强图像缩放造成的目标尺寸特征模糊,对于同类型不同规格特征高度相似零件,往往会发生错判行为。

为此,本文基于深度学习机器视觉检测提出一种零件识别方法,采集样本图片建立尺寸特征模板库,通过使用 PointRend^[10]分割模块,提高原始 Mask R-CNN^[11]实例分割模型分割精度,获得更好分割边界,并基于边界点设计图像尺寸估计算法,获得目标尺寸特征与模板库进行匹配,得到细分结果。本文采集 25 中零件进行识别实验,结果表明,相比较于普通方法,本文方法能够有效提高零件的识别准确性。

1 机械零件识别问题分析与整体方案设计

机械零件包含螺钉、螺母、法兰、垫片等,种类繁多。使用人工识别的方法耗时长、效率低、主观性强;使用传统机器视觉识别的方法场景单一、泛用性差;深度学习方法具有平移不变性、旋转不变性等优点,能够适应多场景、多目标识别,有效替代人工识别,提高检测精度及效率。

图像数据增强是深度学习训练时常用方法,图像缩放是最有效、直接的数据增强手段之一,而缩放会模糊目标尺寸特征。深度学习算法可以有效区分具有明显特征差异的零件类型,但对于特征基本一致,尺寸存在差异的目标,容易出现误判,图 1 为本文设计的机械零件识别方案。

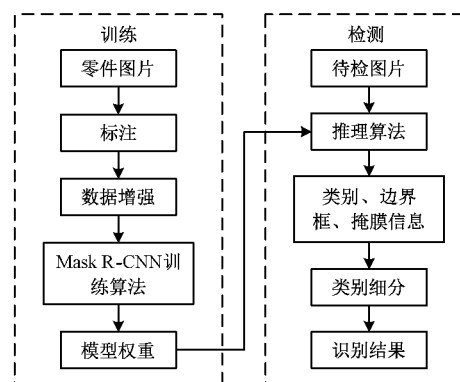


图 1 机械零件识别方案

模型训练时,对同一类型不同规格零件采用相同标签标注,定义大类、小类,大类需要进一步细分,小类无须细分,并仍使用传统数据增强方法。零件检测时,得到目标相关信息,判断是否属于大类;若是大类,则进行类别细分,再得到零件识别结果。

2 基于 Mask R-CNN 的零件识别方法

2.1 Mask R-CNN 图像特征提取

Mask R-CNN 模型架构如图 2 所示,Mask R-CNN 模型是在 Faster R-CNN 模型上添加 Mask 语义分割分支得到的实例分割模型,并使用 RoI Align(regions of interest Align)替代原有的 RoI Pooling 层解决像素对齐问题,能够

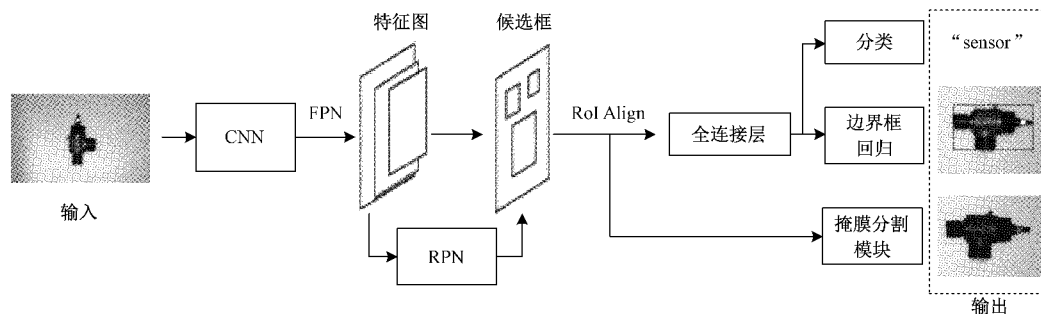


图 2 Mask R-CNN 模型结构

准确进行目标分类、边界框提取与实际像素区域分割,故 Mask R-CNN 可满足本文识别需求。

卷积神经网络(convolutional neural network, CNN)骨干网络对输入图像进行特征提取,并通过特征金字塔网络(feature pyramid networks, FPN)获取一系列特征图;区域推荐网络(region proposal network, RPN)生成一系列的锚点框对特征图进行筛选,得到候选区域;最后通过输入到分类、边界框回归、语义分割模块,得到最终识别结果。

Mask R-CNN 的损失函数定义为分类、回归、分割损失之和:

$$L = L_{cls} + L_{reg} + L_{mask} \quad (1)$$

式中: L_{cls} 、 L_{reg} 是分类、边界框回归损失; L_{mask} 为掩膜分割模块的分割损失。在训练时通过不断梯度下降与反向传播调整卷积核参数,总结目标的一般规律,最终在推理时实现目标的识别与分割。

2.2 基于 PointRend 的掩膜分割强化

Mask R-CNN 采取了全卷积网络(fully convolutional networks, FCN)结构,通过将原始图像不断进行卷积、池化操作,将低密度特征转化为高密度特征,再由高密度特征上采样到与原始图像相同大小,获得对象语义信息,而对象边界轮廓在所有对象像素中占比很小,通过掩膜分割模块多层卷积、池化后的高密度特征再上采样过程中势必会丢失部分轮廓特征信息,造成较大的掩膜边界误差,增加对象实际尺寸估计误差。采用 PointRend 可以有效减少上采样比边界误差,提升边界分割精度。图 3 为 PointRend 原理。

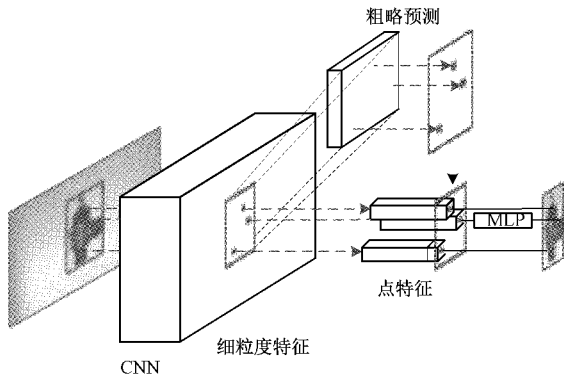


图 3 PointRend 原理

PointRend 将图像渲染思路引进语义分割中,使用双线性插值对 CNN 网络输出的细粒度特征进行上采样,在上采样结果中选取若干个最不确定点(预测概率接近 0.5),作为边界分割困难点,同时结合由 CNN 细粒度特征直接获取的困难点,使用多层感知机(multi-layer perceptron, MLP)进行标签预测。重复上述上采样—选择点—标签预测过程,直到达到所需分辨率。通过在上采样过程中实现自适应细分能够尽可能保留轮廓信息,在分割

中保留边界点,提高分割精度。图 4 为 PointRend 上采样自适应细分过程,图 5 为使用 PointRend 方法的边界分割强化效果。

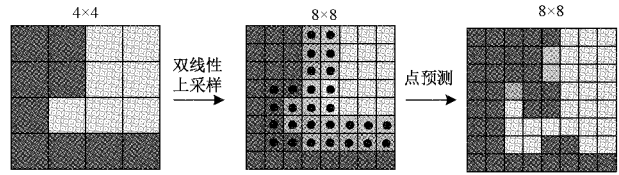
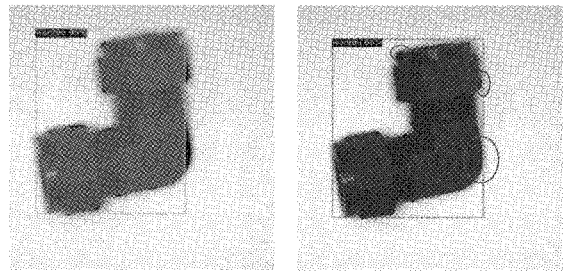


图 4 PointRend 上采样自适应细分过程



(a) Mask R-CNN (b) +PointRend

图 5 基于 PointRend 的边界分割强化效果

2.3 基于深度学习识别结果机械零件细分方法

同一类型的机械零件会存在多个尺寸规格,小样本条件下常采用缩放等方法实现图像数据增强而造成部分图像特征信息丢失,使得深度学习模型难以进行精确区分,必须设计基于深度学习识别结果的机械零件细分方法,将相似样本标注为同一类别,视作大类,同时记录目标像素尺寸,通过尺寸估算实现大类分类细化,得到准确度识别结果。图 6 为机械零件细分过程。

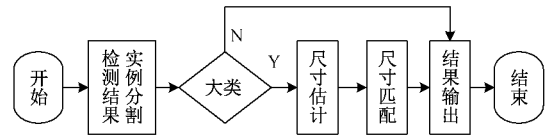


图 6 机械零件细分过程

考虑到实例分割输出的检测信息无法直接获得目标尺寸信息,需对目标轮廓进行最小外接矩形计算,具体流程如下。

1)点集获取。基于深度学习输出目标轮廓点集 P 。使用基于弦到点距离累积(chord-to-point distance accumulation, CPDA)算法^[12]进行角点检测,得到角点集合 P_c ,计算相邻角点间距 d ,若大于设定阈值则在目标轮廓点集 P 。对应区间等距采样 M 点形成点集 P_s 。

2)凸包构建。对离散点集 P_s 使用 Graham 扫描法^[13]得到 m 个凸包点,构建点集 P 。

3)外接矩形计算。对于凸包多边形每条边 $l_i (i = 1, 2, \dots, m - 1)$,计算其与图像水平轴的角度 α ,之后将凸包整体围绕点 $p_i(x_i, y_i)$ 旋转角度 α ,引入齐次表达,则旋转后的点集 P_i 可表示为:

$$P_i = \begin{bmatrix} x_{i1} & y_{i1} & 1 \\ \vdots & \vdots & \vdots \\ x_{im} & y_{im} & 1 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots \\ x_m & y_m & 1 \end{bmatrix} T \quad (2)$$

式(2)中, T 为变换矩阵:

$$T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -x_i & -y_i & 1 \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ x_i & y_i & 1 \end{bmatrix} \quad (3)$$

选取其中 x, y 的极值, 得到 $(x_{\min}^i, y_{\min}^i, x_{\max}^i, y_{\max}^i)$, 组合计算 x, y 极差得到外接矩形的宽高, 并以最小外接矩形宽高作为目标对象的基本像素尺寸。图7为最小外接矩形计算过程。

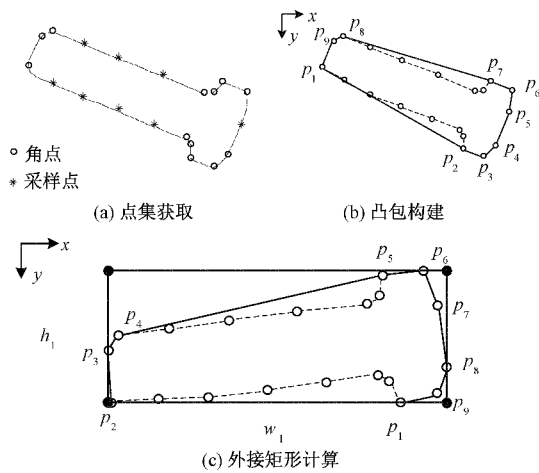


图7 最小外接矩形计算过程

采用小视角镜头, 加大相机到检测平面距离可有效减小由于目标偏离视野中心引起的像素尺寸估计误差。在视野大小为 $0.5 \text{ m} \times 0.34 \text{ m}$ 时, 零件像素尺寸估计误差如表1所示。

表1 零件图像像素尺寸估计误差

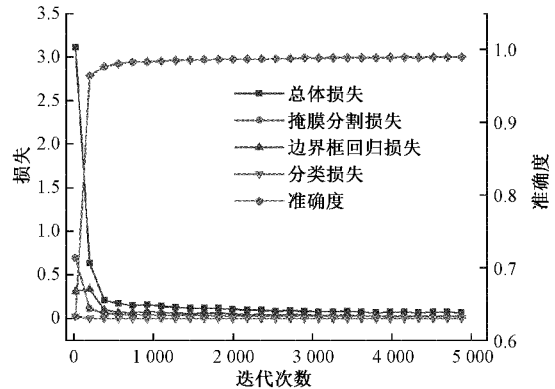
零件	尺寸		估计尺寸		误差/%	
	宽	高	宽	高	宽	高
1	439.0	169.0	450.0	172.0	2.50	1.78
2	865.3	91.4	866.0	89.4	0.08	2.19
3	419.1	214.0	411.0	212.1	1.91	0.89
4	550.1	537.1	564.3	532.7	2.58	0.82
5	170.1	149.1	169.0	147.0	0.65	1.41

通过分析原始数据集标注信息, 通过最小外接矩形算法可得到大类中 J 个细分类别对应零件的标定尺寸 $\{(w_1, h_1), \dots, (w_j, h_j), \dots, (w_j, h_j)\}$, 计算检出尺寸 (w, h) 与标定尺寸间距离, 距离最小时即为其对应细分类别 s :

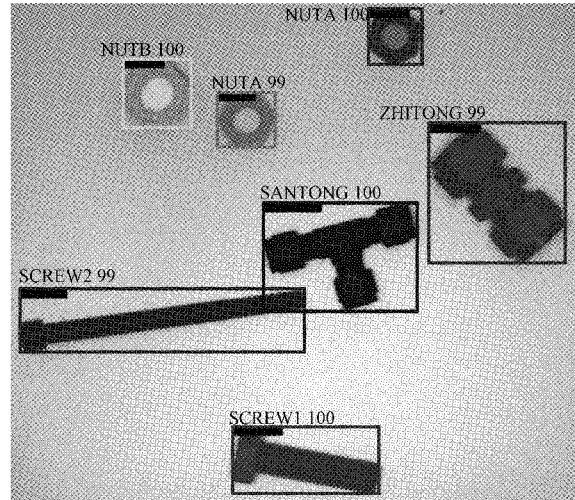
$$s = \operatorname{argmin}_{j=1,2,\dots,J} \sqrt{(w - w_j)^2 + (h - h_j)^2} \quad (4)$$

3 实验与分析

本文共对 25 种机械零件进行检测, 包含非标件 12 种、各规格螺母 5 种、各规格螺栓 8 种。训练集包含图像 240 张, 每种零件均有超过 20 个实例标签, 并使用图像缩放、裁剪、平移等图像增强方法^[14], 扩充数据集大小至 2 400。综合考虑训练时间与训练精度, 使用 ResNet101^[15] 作为骨干网络, 采用 GTX 1080Ti $\times 2$ 配置训练 5 000 次, 获得最终 Mask R-CNN 模型权重, 图8为模型训练结果。



(a) 模型训练参数



(b) 测试集识别结果

图8 模型训练结果

定义识别率 R 、识别准确率 R_{Acc} 为:

$$R = \frac{n_r}{N}, R_{\text{Acc}} = \frac{n_r}{n} \quad (5)$$

其中, N 为验证集中待识别标签数, n 为实际检出标签数, n_r 为检出正确标签数。本文采用相同数据集、相同训练设置训练的原始 Mask R-CNN 作为对照组进行实验, 实验结果如表2所示。可以看出, 本文算法对相似零件识别率达 99.64%, 较原始 Mask R-CNN 方法提升了 3.60%, 识别准确率达 100%, 提升了 11.51%。

通过在 Mask R-CNN 识别基础上进一步细分效果优

表 2 整体识别效果

类型	标签 总数	R/%		R _{Acc} /%	
		对照组	本方法	对照组	本方法
非标件	558	99.81	99.82	100.00	100
螺母	475	95.59	99.55	87.21	100
螺栓	522	92.41	99.63	90.42	100
相似总体	997	93.92	99.55	88.47	100
总体	2 355	96.00	99.64	93.01	100

于 Mask R-CNN 直接识别,能够提高同一类别不同规格零件的识别准确率,在机械零件识别领域效果显著。同时本文算法也可以拓展到其它具有相似特征目标的检测任务中。

4 结 论

本文使用 Mask R-CNN 实例分割模型进行零件检测,引入 PointRend 方法,替代 Mask R-CNN 原始上采样过程,通过自适应细分算法提高了模型的边界分割精度;基于更精确分割边界提出了类别细分方法,重新提取其尺寸特征,有效解决由数据增强中图像缩放导致的相似特征难辨别问题,显著提高同类型不同机械零件识别率。本文实验中,文算法对相似零件识别率达 99.64%,较原始 Mask R-CNN 方法提升了 3.60%,识别准确率达 100%,提升 11.51%。

但本文识别方法仍有局限性,当待区分目标像素尺寸差距不大时,视差影响不可忽略,且针对每个检测场景都需要进行特异化训练。下一步考虑减少视差影响,并结合剪枝对模型进行轻量化改进,进一步提高识别准确度、速度。

参考文献

- [1] 马敬,刘溪,杨勇,等.计算机视觉技术在零件识别中的应用[J].中国设备工程,2020(21):244-246.
- [2] 刘玉珍,朱珍珍,马飞.基于特征融合方法的高光谱图像分类综述[J].激光与光电子学进展,2021,58(4):44-52.
- [3] 吕政阳,邓涛,张丽艳.一种基于机器视觉的飞机钣金件跨粒度识别方法[J].仪器仪表学报,2020,41(2):195-204.
- [4] 魏利胜,丁坤,段志达,等.融合单应性约束 SIFT 特征匹配的轴承滚子检测[J].电子测量与仪器学报,2019,33(9):107-113.

- [5] 石志良,张鹏飞,李晓焱.基于 SIFT-SVM 的发动机主轴轴承盖识别与分类[J].图学学报,2020,41(3):382-389.
- [6] 郑太雄,江明哲,冯明驰.基于视觉的采摘机器人目标识别与定位方法研究综述[J].仪器仪表学报,2021,42(9):28-51.
- [7] 张鹏飞,石志良,李晓焱,等.基于深度学习的主轴承盖分类识别算法[J].图学学报,2021,42(4):572-580.
- [8] 陈丹,林清泉.基于级联式 Faster RCNN 的三维目标最优抓取方法研究[J].仪器仪表学报,2019,40(4):229-237.
- [9] 杨琳,陈赛旋,崔国华,等.基于改进 YOLOv4 算法的零件识别与定位[J].组合机床与自动化加工技术,2021(10):28-32,37.
- [10] KIRILLOV A, WU Y, HE K, et al. Pointrend: Image segmentation as rendering[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 9799-9808.
- [11] HE K, GKIOXARI G, DOLLÁR P, et al. Mask r-cnn [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [12] 白萌萌,景军锋,章为川,等.边缘轮廓上基于 CPDA 的图像角点检测[J].西安工程大学学报,2021,35(2):29-34.
- [13] 孙聪,初建宇,李印凤,等.基于坐标排序 Graham-scan 的飞行受限区划设方法[J].指挥信息系统与技术,2020,11(5):40-44.
- [14] 吴睿曦,肖秦琨.基于深度网络和数据增强的多物体图像识别[J].国外电子测量技术,2019,38(5):86-90.
- [15] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

作者简介

臧春华,高级工程师,主要研究方向为计量测试、仪器仪表。

E-mail:16850433@qq.com

周介祺,硕士研究生,主要研究方向为现代传感技术与机器视觉。

E-mail:zjq0796@foxmail.com

刘桂雄,教授,主要研究方向为先进传感与仪器研究。

E-mail:megxliu@scut.edu.cn