

# 基于深度学习的元器件视觉识别和定位技术\*

雷文桐 顾寄南 胡君杰 高伟  
(江苏大学机械工程学院 镇江 212013)

**摘要:**为解决当前装配机器人视觉系统对元器件误检率高、效率低、难获取有效定位信息的问题,提出了一种基于深度学习的元器件视觉识别和定位方法。首先,设计基于深度聚合和解耦头的高精度检测算法,提高元器件识别和主体检测的精度;其次,设计标注和判定规则,细化定位主体轮廓和抓取点;最后,设计基于网络剪枝的轻量化检测算法,实现模型压缩,提高引脚检测和装配点定位的效率。研究表明:该方法在元器件的识别和定位上取得了较好表现,类别识别平均错误率仅为0.27%,计算量减少了29.8%,参数量减少了22.7%,并将传统的元器件轮廓检测扩展到抓取点和装配点定位,得到丰富的类别和位置指引信息,为工业机器人精准、可靠、稳定地抓取和装配做好基础。

**关键词:** 目标检测;智能装配;电子元器件;机器视觉;深度学习

**中图分类号:** TP391;TH166 **文献标识码:** A **国家标准学科分类代码:** 520.2060

## Component vision recognition and location technology based on deep learning

Lei Wentong Gu Jinan Hu Junjie Gao Wei  
(School of Mechanical Engineering, Jiangsu University, Zhenjiang 212013, China)

**Abstract:** In order to solve the current assembly robot vision system's problems of high false detection rate, low efficiency, and difficulty in obtaining effective positioning information, a component vision recognition and positioning method based on deep learning was proposed. Firstly, a high-precision detection algorithm based on deep aggregation and decoupling head was designed to improve the accuracy of component identification and subject detection. Secondly, the rules of labeling and determination were designed, and the position subject outlines and grasping points were refined. Lastly, a lightweight detection algorithm based on network pruning was designed to accomplish model compression and improve the efficiency of pin detection and assembly point positioning. The research results show that the method has achieved better performance in the identification and positioning of components. The average error rate of category recognition is merely 0.27%. The calculation is reduced by 29.8%, and the volume of parameters decreased by 22.7%. Through this method, traditional component contour detection is extended to grasp point and assembly point positioning to obtain abundant category and position guideline information, laying a foundation for industrial robots to grasp and assemble accurately, reliably and stably.

**Keywords:** target detection; intelligent assembly; electronic components; machine vision; deep learning

## 0 引言

电子元器件是电子电路中的基本元素,需装配到印制电路板(printed circuit board, PCB)上实现相应功能<sup>[1]</sup>。但当前元器件的装配过程多依靠人工完成,工人先在料箱或传送带上取所需元器件,插在PCB板上对应位置。这要求工人能够准确且高效地判别元器件类别,但长时间的工作,容易造成工人的视觉疲劳,导致元器件的安装失误<sup>[1]</sup>。

随着智能制造的发展,机器视觉技术作为一种新兴的技术,通过机器代替人眼做测量和判断,得到制造过程所需要的信息,目前在自动驾驶、人脸识别、产品检测等领域得到广泛应用<sup>[2]</sup>。装配场景下利用机器视觉技术进行元器件识别和定位,能为工业机器人灵活、可靠、稳定抓取和装配提供基础。

目标检测是机器视觉的一个分支,分为传统方法和深度学习<sup>[3-4]</sup>。传统方法多利用手动设计得到物体特征,

收稿日期:2022-11-04

\* 基金项目:国家自然科学基金(51875266)项目资助

流程包括:区域选择→特征提取→分类器<sup>[4]</sup>。代表性方法有滑动窗口法、SIFT 法、HOG 法<sup>[3-4]</sup>,但计算量大、设计冗余、在密集网格上低效,泛化能力差。深度学习方法利用卷积神经网络自主提取物体特征。2015 年 Zaidi 等<sup>[5]</sup>提出的 Faster RCNN 是第一个端到端的深度学习模型,但采用的区域建议网络计算繁重。单阶段算法在检测速度上取得显著进步。2016 年 Liu 等<sup>[6]</sup>提出的 SSD 采用多分辨率技术提高小目标的检测精度。2017 年 Lin 等<sup>[7]</sup>提出的 RetinaNet 引入了 focal loss,让单阶段算法实现了较高精度,但锚框设计冗余<sup>[8]</sup>。旋转目标检测在通用目标检测的基础上加入了角度信息。2021 年 Yang 等<sup>[9]</sup>提出了 R3Det,2022 年 Hou 等<sup>[10]</sup>提出了旋转的 YOLO 算法 R-YOLOv5,在轮廓框的准确包围上取得了突破性的进展。

总体来说,传统目标检测具有设计冗余、低效、泛化能力差的缺点。当前流行的深度学习目标检测的缺点是将重点放在目标物类别的判断上,在边界框定位上做的不够好。最突出的缺陷是得到的包围框是外接矩形,不能契合真实场景中多角度多姿态的物体,且对密集物体容易产生漏检、误检的情况<sup>[11-12]</sup>。旋转目标检测的缺点是多用于遥感场景中尺寸相近且方位统一的船舶、车辆检测,对于复杂装配场景中的元器件检测效果较差<sup>[13]</sup>,需针对元器件多尺度、排布无序的特性改进算法。

因此,建立了一个面向装配工位单元群电子元器件智能装配移动机器人的视觉识别和定位模型。具体包括三部分:主体和引脚检测算法、抓取点定位算法、装配点定位算法,系统地降低了元器件误检率、提升了检测效率、提供了更丰富的定位信息。

### 1 基于深度聚合和解耦头的高精度检测算法

#### 1.1 高精度网络 DGDNet v1

近年来随着机器视觉任务的广泛应用和不断演进,对高性能算法的追寻也在持续盛极。在提升卷积神经网络的性能过程中,主干网的设计成了一个非常重要的议题。本文设计了一个高精度网络 DGDNet v1,创新点在于:首先,通过两种跳跃连接策略融合语义和空间特征。其次,采用 Ghost 模块解决了特征图冗余的问题。最后,采用可变性卷积 DCN 有效地适应目标物体的几何外观变化。

##### 1) 深度聚合跳跃连接策略 DLA

为了拥有更好的性能,主干网络层数不断加深。但一味增加网络深度会导致梯度发散和模型退化,导致模型难收敛<sup>[14]</sup>。为解决此问题,近年提出了捷径分支结构,如图 1 所示,但此结构单一、无针对性,不能针对目标物差异调整融合语义和空间特征的策略。

针对此问题,本文采用两种跳跃连接策略<sup>[15]</sup>:迭代深度聚合策略(iterative deep aggregation, IDA)、分层深度聚合策略(hierarchical deep aggregation, HDA),改善当前流行捷径分支,得到更优的网络特征融合效果。

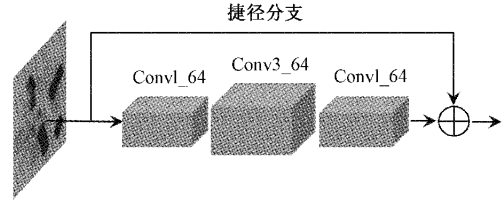


图 1 捷径分支结构

迭代深度聚合结构(IDA)如图 2 所示,致力于各阶段外部的聚合。通过在各阶段外部增加捷径分支和聚合节点,在网络由浅入深传播的过程中聚集特征。

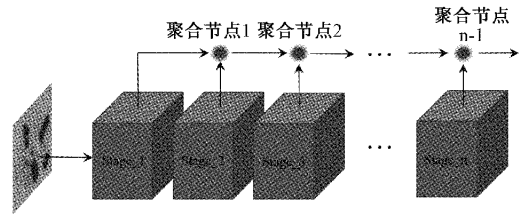


图 2 迭代深度聚合结构 (IDA)

分层深度聚合结构(HDA)如图 3 所示,致力于各阶段内部的块的聚合,将不同特征进行交流,提高模型对不同特征物体识别的鲁棒性<sup>[15-16]</sup>。

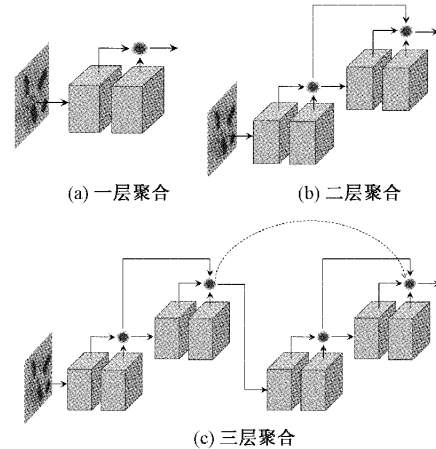


图 3 分层深度聚合结构 (HDA)

将 IDA 和 HDA 共享聚合节点,得到了迭代分层深度聚合,简称深度聚合(DLA),结构如图 4 所示。

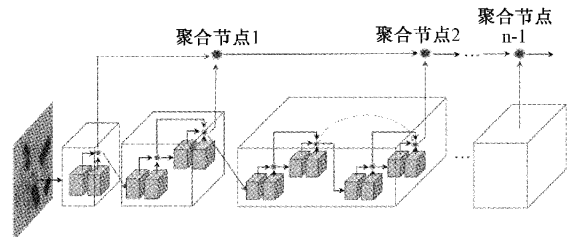


图 4 迭代分层深度聚合结构 (DLA)

以 DLA 为基础构建基准网络 DLA6Net34。主干网包括 6 个阶段,各阶段间会进行降采样操作,选用卷积核大小

为2,步长为2的最大池化进行降采样。各层结构参数如 表1所示,整体会对输入图像进行32倍降采样。

表1 DLA6Net34 网络结构

网络层数	输入	通道数	聚合深度	模块类型	输出
0(输入)	$1 \times 512 \times 512 \times 3$	3	/	Sequential[base_layer]	$1 \times 512 \times 512 \times 16$
1	$1 \times 512 \times 512 \times 16$	16	/	Sequential[level0]	$1 \times 256 \times 256 \times 32$
2	$1 \times 256 \times 256 \times 32$	32	/	Sequential[level1]	$1 \times 128 \times 128 \times 64$
3	$1 \times 128 \times 128 \times 64$	64	1	Tree[level2]	$1 \times 64 \times 64 \times 128$
4	$1 \times 64 \times 64 \times 128$	128	2	Tree[level3]	$1 \times 32 \times 32 \times 256$
5	$1 \times 32 \times 32 \times 256$	256	2	Tree[level4]	$1 \times 16 \times 16 \times 512$
6	$1 \times 16 \times 16 \times 512$	512	1	Tree[level5]	$1 \times 16 \times 16 \times 512$
7	$1 \times 16 \times 16 \times 512$	512	/	AvgPool2d[avgpool]	$1 \times 2 \times 2 \times 16$
8	$1 \times 2 \times 2 \times 16$	16	/	Conv2d[fc]	$1 \times 2 \times 2 \times 16$
9(输出)	$1 \times 2 \times 2 \times 16$	/	/	View	$1 \times 64$

## 2) Ghost 模块

为讨论特征图冗余问题,对特征图进行可视化,如图5所示,相同颜色框出的特征图相似度高,将此类相似性高的图称为彼此的 Ghost 冗余特征图<sup>[17-18]</sup>。经计算,Ghost 冗余特征图占第二阶段特征图的18.7%,后续阶段随语义信息减少,比例升高。本文使用廉价转换 Ghost 模块代替  $3 \times 3$  卷积操作生成特征图,降低模型参数。

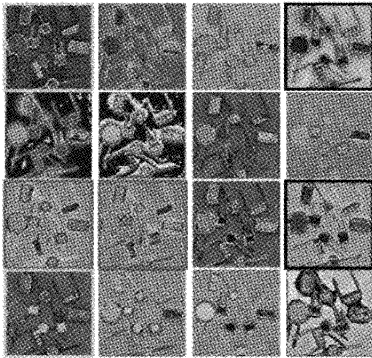
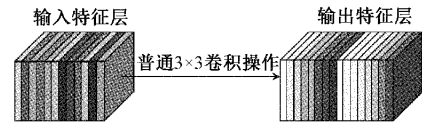


图5 DLA6Net34 第二层特征图的可视化结果

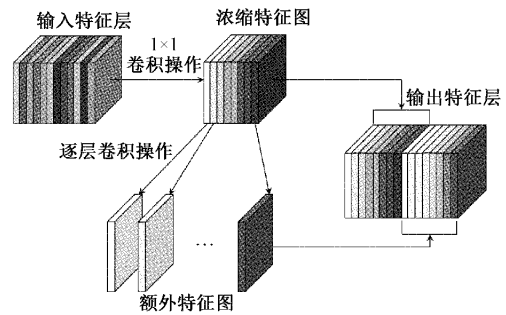
Ghost 模块结构如图6所示,分为特征浓缩和逐层卷积两步<sup>[17-18]</sup>。首先,利用1个  $1 \times 1$  卷积实现通道数压缩;其次,利用逐层卷积进行跨特征点特征提取,生成相似的额外特征图;最后,将浓缩特征图与额外特征图进行堆叠,获得输出特征层。原理本质是将一个卷积乘法拆解为两个卷积乘法的相加,从而达到减少参数数量的目的。

## 3) 可变形卷积 DCN

固定形状的卷积核对形状大小统一的物体尚能较好提取特征,但对形状各异且未知的物体存在天生的弱势<sup>[18]</sup>。可变形卷积(deformable convolutional networks, DCN)在各采样点上添加偏移量,能提升卷积对目标物体几何变化的适应能力<sup>[19]</sup>。本文在特征提取的中后部分用 DCN 代替普通卷积,提高了拟合精度。



(a) 普通  $3 \times 3$  卷积操作



(b) Ghost模块

图6 普通卷积操作和 Ghost 模块的结构对比

$$F(p_0) = \sum_{p_n \in \lambda} Z(p_n) \cdot X(p_0 + p_n) \quad (1)$$

$$F(p_0) = \sum_{p_n \in \lambda} Z(p_n) \cdot X(p_0 + p_n + \Delta p_n) \quad (2)$$

其中,  $\lambda$  为输入特征图的规则采样区域。式(1)表示普通卷积的特征提取过程,对采样值和权重  $X$  求和,得到特征点  $p_0$  的输出特征。式(2)表示 DCN 的特征提取过程,对每个特征点进行偏移  $\Delta p_n$ <sup>[19]</sup>。

$$X(p) = \sum_q G(q, p) \cdot X(q) \quad (3)$$

$$G(q, p) = g(q_x, p_x) \cdot g(q_y, p_y) \quad (4)$$

数学上常用式(3)、(4)的双线性插值计算偏移位置,其中  $G(q, p)$  为双线性插值权重。

## 1.2 解耦合检测头

### 1) 常规检测头

全连接头和卷积头是深度学习领域两个经典的检测头。其中,全连接头与分类任务匹配度较强,更适合分类任务,卷积头对空间敏感,能提供精确的边界框回归,更适

合定位任务<sup>[8]</sup>。

2) 解耦合检测头

由于分类更关注纹理特征,定位更关注边缘信息,合并到同一分支输出会造成信息混乱且缺乏针对性<sup>[3-4]</sup>,故常规检测头在某些视觉任务上不是最佳选择。

解耦合检测头通过几个不同的分支解耦出类别、位置信息,虽耗费少许额外成本,但能大大加快各分支的收敛速度,提高分类和定位精度<sup>[18]</sup>。

3) 探究角度分支最优位置

原始解耦合检测头由中心点分支、尺寸分支组成。本文改进检测头,添加角度分支。为了探究角度分支的最优位置,分别将角度分支合并到中心点分支和尺寸分支上,观察各损失收敛情况,结果如图 7 所示。

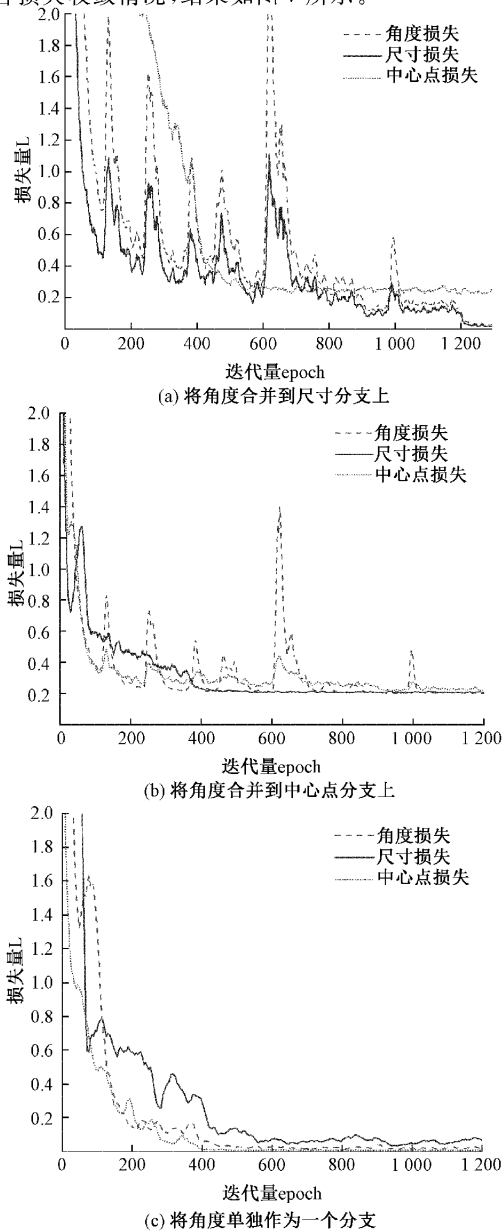


图 7 角度分支在不同位置时的网络训练结果对比

图 7(a)将角度合并到尺寸分支上,角度损失和尺寸损失波动明显且相互干扰,最终未收敛,中心点分支最终收敛;图 7(b)表示角度合并到中心点分支上,角度损失和中心点损失仍相互干扰,虽波动较(a)平稳,但最终仍未收敛;图 7(c)表示角度单独作为一个分支,3 个分支在 600 次迭代时均已收敛,可见训练效果较好。

笔者认为,角度和尺寸、中心点的共享特征不明确,且会相互干扰,合并到同一分支上难以分配和表征不同的信息,所以很难收敛。因此,选择将角度分支作为解耦头的第三个单独分支。

2 元器件抓取点定位算法

2.1 带角度标注

目标检测数据集的标注工具是 LabelImg,但本研究需要带角度框的标注,故采用改进的可视化图像标定工具 LabelImg2,在中心点  $(x, y)$ 、宽  $w$ 、高  $h$  的 4 个变量的基础上加入角度信息  $\alpha$ ,标定工具软件如图 8 所示。

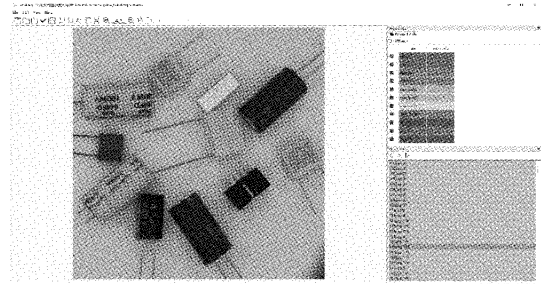


图 8 LabelImg2 图像标定工具软件界面

针对元器件特性,定义标注准则:以初始标注框内中心点向上延伸的指引虚线为基准,将标注框按顺时针旋转至指引虚线与引脚朝向一致。各图片标注完成会生成 PASCAL VOC 格式的 .xml 文件,包含图片中各元器件的类别和位置信息。

2.2 对边和侧边的判定

在带角度标注方法和自定义的标注准则的基础上,进一步区别元器件对边和侧边。其中,本文规定元器件对边指检测框四条边中元器件的引脚位置的两条边,元器件侧边指引脚位置相邻的两条边。

$$y_{N_n} = (x_n - x) \cdot \sin\alpha + (y_n - y) \cdot \cos\alpha + y \quad (5)$$

$$x_{N_n} = (x_n - x) \cdot \cos\alpha - (y_n - y) \cdot \sin\alpha + x \quad (6)$$

式(5)、(6)指加入角度的最小外接矩形的四点坐标。

其中,  $n$  代表矩形框的四点,  $n \in [0, 1, 2, \dots]$ ,  $N$  代表一张图像中的各元器件,  $N \in [A, B, C, \dots]$ 。

本文规定  $N_0(x_{N_0}, y_{N_0})$  为引脚方向左上角点,  $N_1, N_2, N_3$  依次逆时针循环,如图 9。因此,  $N_0N_3, N_1N_2$  为元器件的两条对边,用红色边框。  $N_0N_1, N_2N_3$  为元器件的两条侧边,用蓝色边框表示,如图 10。

2.3 抓取点定位

元器件主体部分原料通常是金属(金、银、铜、铁)、非

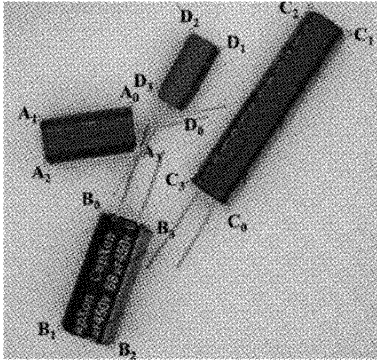


图9 检测框四点顺序

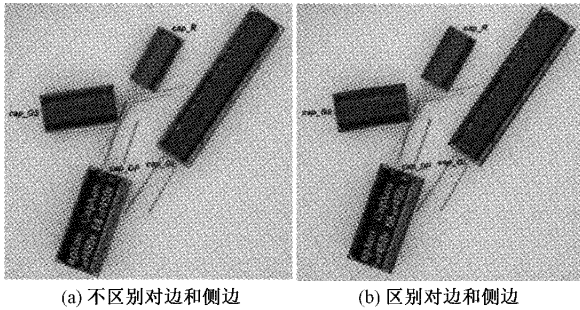


图10 检测结果对比

金属(碳)以及半导体(硅、锗、磷砷化镓),元器件引脚部分原料通常为铜铁材料,两者密度不相上下<sup>[1]</sup>。鉴于元器件主体部分体积远远大于引脚,结合机器人抓取原理,故将元器件中心近似看作元器件重心。

$$\begin{aligned} P_{Grab1} &= \left( \frac{x_{N0} + x_{N1}}{2}, \frac{y_{N0} + y_{N1}}{2} \right) \\ P_{Grab2} &= \left( \frac{x_{N3} + x_{N2}}{2}, \frac{y_{N3} + y_{N2}}{2} \right) \end{aligned} \quad (7)$$

如式(7)所示,对于机械臂末端夹爪的抓取任务来说,夹爪与元器件垂直接触,且两个接触点的中心是重心,如此能够实现稳定可靠的抓取。

### 3 基于网络剪枝的引脚轻量化检测及装配点定位算法

为得到元器件的完整信息,需进一步对元器件引脚进行识别和定位。各元器件和对应引脚具有从属关系,故引脚检测有自底向上和自顶而下两个思路。自底向上指先检测出各元器件和各引脚,再进行匹配。自顶而下指先检测元器件,再对检测到的元器件进行引脚检测。

由于一张图像中有多个元器件,自顶向下法会将截取后的元器件依次检测,效率不如自底向上法,但自底向上法匹配错误率太高。为了实现高精度定位,选用采用自顶向下法,并通过轻量化操作实现效率提升。本文设计了一个轻量化网络 DGDNet v2,创新点在于:通过权重排序和阈值划分进行关键点通道和部分 BN 层的剪枝,精准地舍

弃效果差的层结构。

#### 3.1 单个元器件截选

自顶向下法需要在元器件检测的基础上进行单个元器件图片的截选。截选流程如下:

1)将包围框向引脚方向按比例延伸,延伸比例如表 2。

表2 各元器件延伸比例

元器件类型	延伸比例
热敏电阻	2.5 倍
色环电阻	6 倍
二极管	3 倍(两端)
安规电容	2.5 倍
电解电容	0.7 倍(长宽比 $\geq 3$ ) 1.5 倍(长宽比 $\leq 3$ )
插座端子	1 倍

2)在结果图中隐藏原始的旋转包围框,显示延伸后的包围框,保存到文件夹中。

3)利用 opencv 的 cv2. boundingRect 函数得到各外接矩形的中心点坐标和宽高。

4)利用 opencv 的 cv2. rectangle 函数得到各外接矩形。

5)利用各中心点和宽高进行切片操作,将每个外接矩形进行裁剪、保存。将裁剪后的外接矩形左上角点记为  $N_i$ 。

#### 3.2 轻量化网络 DGDNet v2

将裁剪后的图片用 Labelimg2 进行标注。由于每张图片中目标物数量少且类别为 1 类,检测难度低于本体的检测,进行网络剪枝和轻量化得到 DGDNet v2。

1)关键点通道剪枝

在定位任务下只需要输出引脚的包围框,对于关键点没有需求,故剪枝关键点通道<sup>[18]</sup>。在精度几乎保持不变情况下,计算量减少了 24.69%,参数量减少了 14.57%。

2)BN 层剪枝

首先,利用 BN 层中 gamma 系数评价各通道的所占权重,将权重进行排序<sup>[13]</sup>。其次,设定阈值对部分权重值赋零,剪掉 BN 对应的低权重通道<sup>[5]</sup>。接着,变换阈值,并计算各阈值对应的 ap 值变换曲线,得到 ap 值稳定时的最大阈值<sup>[5]</sup>。最后,根据剪枝后的 config 重构模型。

#### 3.3 引脚装配点定位

引脚装配点为各引脚末端的中点,通过 opencv 进行引脚装配点定位。装配点坐标如式(8)所示。

$$P'_{Assemble} = \left( \frac{(x'_{N0} + x'_{N1})}{2}, \frac{(y'_{N0} + y'_{N1})}{2} \right) \quad (8)$$

此坐标为装配点相对于裁剪后各外接矩形左上角  $N_i$  的相对坐标,其中  $N \in [A, B, C \dots]$ 。将相对坐标通过仿射变换转化为裁剪前原图中的绝对坐标,如式(9)所示。

$$P_{Assemble} = \left( x_{N1} + \frac{(x'_{N0} + x'_{N1})}{2}, y_{N1} + \frac{(y'_{N0} + y'_{N1})}{2} \right) \quad (9)$$

### 4 实验及结果分析

#### 4.1 实验环境

##### 1) 图像采集实验平台

图像采集实验平台由 DEMO 球形光源、工业相机、相机支架、显示器、工控机组成,如图 11 所示。选用 CMOS 相机采集图像。选用 VRD70 型号的 DEMO 球形光源进行打光,其外径 700 mm、通光孔径 120 mm、24 V/8 W。

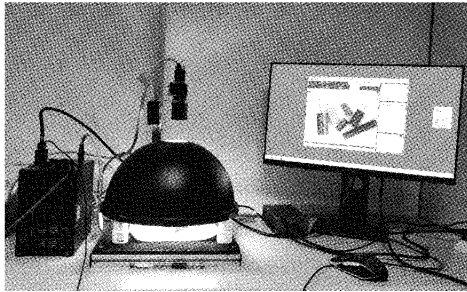


图 11 图像采集实验平台

##### 2) 实验环境与配置

实验环境与配置如表 3 所示。在训练前会将输入图像分辨率统一转成  $512 \times 512$ 。训练中图片批量大小为 10,进行 300 次迭代循环。

表 3 实验环境与配置

开发环境	型号/参数
GPU 型号	NVIDIA RTX2080Ti
CPU 型号	Intel Xeon(R) Gold 5118
操作系统	Ubuntu18.04(64 bit)
集成开发环境	PyCharm Community Edition(64 bit)
编程语言	Python 3.7 (64 bit)
深度学习框架	PyTorch1.7.1
训练优化器	Adam 优化器
学习率	$1.25 \times 10^{-4}$
权重衰减率	$1 \times 10^{-4}$

#### 4.2 数据集构建

当前电子元器件检测领域没有公共数据集,本文自行构建两个电子元器件数据集。

##### 1) 元器件主体数据集

元器件主体数据集包括 20 类元器件,如图 12 所示。采集 500 张图像,用标注工具 LabelImg2 生成 .xml 文件。通过图像增强提高样本多样性,将图像扩充到 5 000 张。将 5 000 个 .voc 格式的 .xml 文件转化为 coco 格式的 .json 文件,按照 8 : 2 的比例划分训练集和验证集。

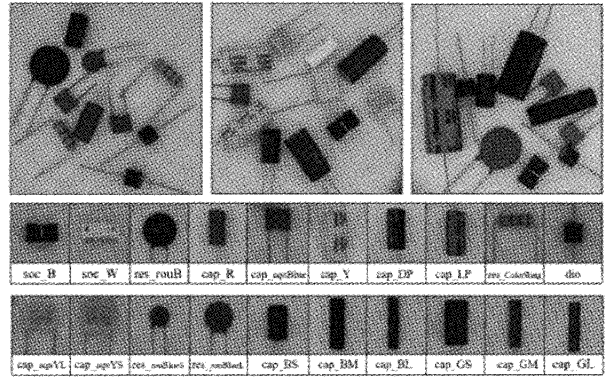


图 12 元器件主体数据集

##### 2) 元器件引脚数据集

元器件引脚数据集包括 1 类引脚,350 张图像由元器件检测后进行单个元器件筛选得到,通过数据增强扩充至 1 050 张。将 1 050 个 .voc 格式的 .xml 文件转化为 coco 格式的 .json 文件,按照 7 : 3 的比例划分训练集和验证集。两个数据集的划分结果如表 4 所示。

表 4 数据集划分结果

数据集类型	元器件主体数据集	元器件引脚数据集
原始图像总数	500 张	350 张
增强后图像总数	5 000 张	1 050 张
数据集类别数	20 类	1 类
标注目标总数	39 620 个	2 625 个
训练集数目	4 000 张	735 张
验证集数目	1 000 张	315 张

#### 4.3 评价指标

目标检测领域常用精度和效率评价算法性能<sup>[18]</sup>。本文用  $F_1$  和  $mAP$  衡量模型精度,用计算量  $FLOPs$ 、参数量  $params$ 、每秒检测帧数衡量模型效率。

##### 1) 精度评价指标

对于类别  $k$ ,  $f_{1k}$  是准确率与召回率的调和平均数,  $F_1$  为所有类别的平均值<sup>[4]</sup>,如式(10)、(11):

$$f_{1k} = 2 \frac{Recall_k \times Precision_k}{Recall_k + Precision_k} \quad (10)$$

$$F_1 = \left( \frac{1}{n} \sum f_{1k} \right)^2 \quad (11)$$

$AP$  为单类别的差值平均准确率,  $mAP$  是各类别  $AP$  的平均值<sup>[4]</sup>,如式(12)(13):

$$AP = \sum_{i=1}^{m-1} (r_{i+1} - r_i) Pinterp(r_{i+1}) \quad (12)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (13)$$

##### 2) 效率评价指标

参数量  $params$  是衡量算法空间复杂度的指标<sup>[18]</sup>。卷

积层参数量如式(14),其中  $C_{in}$ 、 $C_{out}$  为输入和输出通道数,  $K_w$ 、 $K_h$  为卷积核宽高。

$$params_{conv} = (C_{in} \times K_w \times K_h + 1) \times C_{out} \quad (14)$$

计算量  $FLOPs$  是衡量算法时间复杂度的指标<sup>[18]</sup>,卷积层的计算量如式(15):

$$FLOPs_{conv} = 2 \times C_{in} \times K_w \times K_h \times W \times H \times C_{out} \quad (15)$$

#### 4.4 实验结果

##### 1) 主体检测算法 DGDNet v1 实验验证

DLA6Net34 与流行主干网的性能对比如表 5 所示。与深层主干网 DenseNet 相比,DLA6Net34 计算量和参数量分别降低了 69.5 G 和 37.258 M;与浅层残差 ResNet18、ResNet34 和轻量化网络 MobileNetv3 相比,DLA6Net34 的  $mAP$  分别增加 7.04%、5.45%、3.46%,每秒检测帧数分别增加了 25.6 帧、30.4 帧、13.3 帧,这表明 DLA6Net34 的准确率和效率均大幅提高。

表 5 主干网性能对比

主干网络	F1	mAP	每秒检测帧数	计算量/参数量/ G M	
				G	M
ResNet18	0.874	0.867	32.5	72.965	17.734
ResNet34	0.859	0.880	27.7	79.354	27.842
DenseNet	0.946	0.945	16.9	97.958	55.349
MobileNetv3	0.905	0.897	44.8	85.158	36.359
DLANet60	0.957	0.933	36.1	60.336	31.367
DLA6Net34(ours)	0.937	0.928	58.1	28.458	18.091

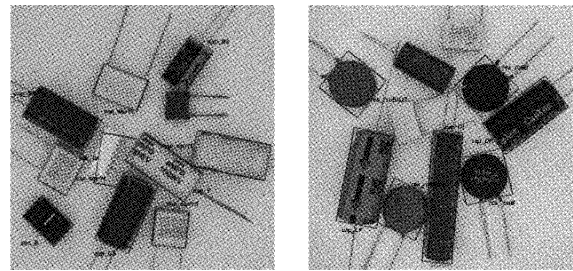
将 DLA6Net34 作为基准模型,通过消融实验验证 Ghost 和 DCN 的贡献。由表 6 可知,单独增加 Ghost,  $mAP$  提升 0.22%,证明 Ghost 模块生成特征图的高性能。单独增加 DCN,  $mAP$  提升 0.75%,证明 DCN 提高了拟合精度。将 Ghost 和 DCN 一同添加,在牺牲少许效率的情况下  $F_1$  提升 2.13%,  $mAP$  提升 1.4%,验证了 DGDNetv1 的准确性。

表 6 DGDNetv1 消融实验结果

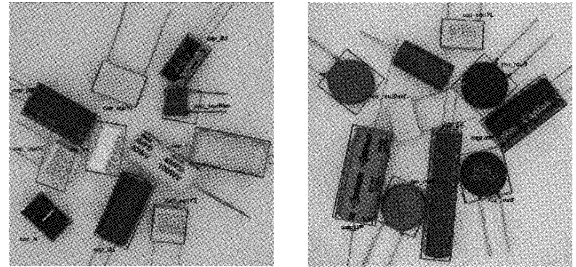
模型	改进方法		评价指标			
	Ghost	DCN	F1	mAP	计算量/参数量/ G M	
DLA6Net34	×	×	0.937	0.928	28.458	18.091
	√	×	0.948	0.930	24.983	16.637
	×	√	0.952	0.937	32.768	22.654
DGDNetv1	√	√	<b>0.957</b>	<b>0.941</b>	29.295	21.201

为更直观地展示检测效果,将当前流行的旋转目标检测 R3Det 和精度较高的 R-YOLOv5 进行检测效果可视化

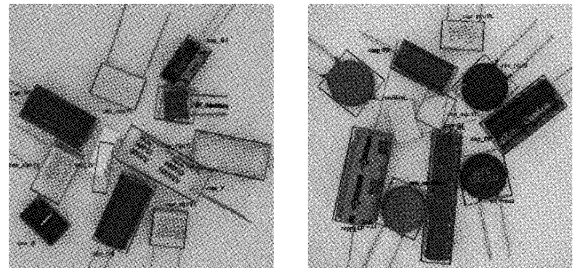
对比。如图 13 所示,对比模型在分类或定位方面有以下错误:DLA6Net34 漏检蓝色方形电容,将右侧 cap\_Y 误检为 cap\_squYL。R3Det 在类别上未出错,但漏检黄色电容。R-YOLOv5 未产生漏检,但检测框未紧密包围目标物,且对 cap\_squBlue 二次重复检测。分类错误的原因是 DLA6Net34 层数较少,特征提取性能不足。定位错误的原因是 R3Det 和 R-YOLOv5 引入大量的旋转 Anchor,并采用松弛匹配策略,导致正样本的错位。DGDNetv1 在类别识别和包围框紧密包围上都表现较好,无漏检和误检,具有较高的精确度。



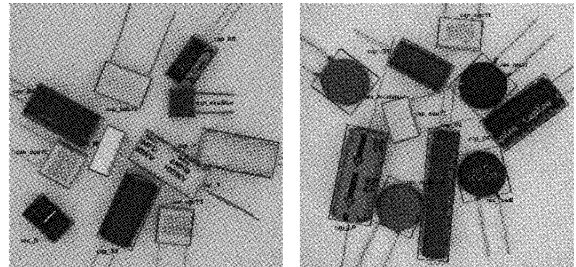
(a) DLA6Net34模型



(b) R3Det模型



(c) R-YOLOv5模型



(d) DGDNetv1模型(本文)

图 13 元器件主体检测结果

随机抽取一批未训练图片进行类别错误率测试。由表 7 可知,R-YOLOv5 错误率为 9.65%,R3Det 为 6.86%,

而 DGDNetv1 各类别平均错误率仅为 0.27%，由此证明了 DGDNetv1 在类别的准确判断上具有极大优势。

表 7 各类别检测错误

类别/算法	R-YOLOv5	R3Det	DGDNet v1 (本文)
soc_B	0.03	0.120	0
soc_W	0	0	0
res_rouB	0.027	0.054	0.012
cap_R	0	0.004	0
cap_squBlue	0.020	0.034	0
cap_Y	0.140	0.252	0
cap_DP	0	0.008	0
cap_LP	0	0	0
res_ColorRing	0	0	0
dio	0.110	0	0
cap_squYL	0.236	0.213	0.020
cap_squYS	0.240	0.117	0.014
res_rouBlueS	0.230	0.179	0
res_rouBlueL	0.417	0.314	0.008
cap_BS	0	0.007	0
cap_BM	0	0	0
cap_BL	0.210	0	0
cap_GS	0.008	0	0
cap_GM	0.120	0.047	0
cap_GL	0.142	0.023	0
平均错误率	9.65%	6.86%	<b>0.27%</b>

2) 引脚检测 DGDNet v2 实验验证

同理,将 DLA6Net34 作为基准模型,通过消融实验验证通道剪枝和 BN 层剪枝对算法轻量化的贡献。从表 8 可以看出,单独进行通道剪枝,在  $F_1$  和  $mAP$  牺牲了少许精度,但计算量和参数量大幅减少,由此证明了通道剪枝能够有效轻量化网络。单独进行 BN 层剪枝,计算量减少了 12.9%,参数量减少了 6.35%,由此证明 BN 层剪枝能够实现模型压缩。对 DLA6Net34 同时进行通道和 BN 层剪枝,在牺牲少许精度的情况下计算量减少了 29.8%,参数量减少了 22.7%,验证了 DGDNetv2 的高效性。

表 8 DGDNetv2 消融实验结果

模型	改进方法		评价指标			
	通道剪枝	BN 层剪枝	F1	mAP	计算量/ G	参数量/ M
DLA6Net34	×	×	0.937	0.928	28.458	18.091
	√	×	0.932	0.924	21.435	15.454
	×	√	0.924	0.922	24.765	16.943
DGDNetv2	√	√	0.919	0.917	<b>19.987</b>	<b>13.987</b>

为更直观地展示检测效果,进行检测效果的可视化对比。如图 14 所示,R3Det 得到的检测框角度偏差较大,检测框和目标物的位置拟合精度较低,这是由于 R3Det 是针对遥感领域小目标的检测,在推理过程会自动裁剪图像,在各个小图像上进行推理再合并,这不适用于本研究场景。DGDNet v2 在检测框的拟合上表现较好。

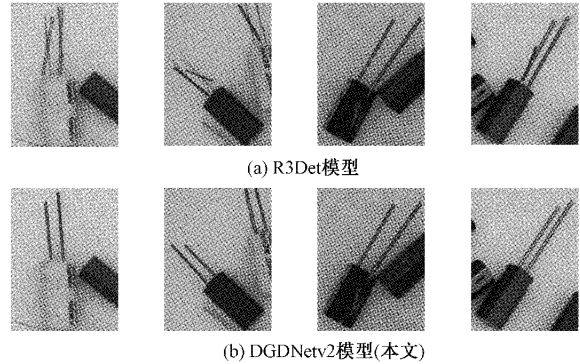


图 14 元器件引脚检测结果

3) 抓取点和装配点定位实验验证

利用 DGDNet v1 检测元器件主体,利用 DGDNet v2 检测元器件引脚,进一步进行主体抓取点和引脚装配点定位。本研究最终总算法检测结果如图 15 所示。

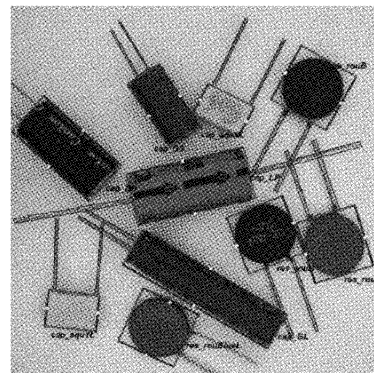


图 15 元器件检测及抓取点和装配点定位结果

5 结 论

本研究主要有以下贡献:

针对装配机器人识别元器件的误检率高、效率低的问题,构建了两个元器件数据集,提出了高精度算法 DGDNetv1、轻量化算法 DGDNetv2。最终检测类别平均错误率仅为 0.27%,计算量减少了 29.8%,参数量减少了 22.7%,提高了元器件识别和检测的准确率和效率。

针对元器件装配过程缺乏有效定位信息的问题,设定标注和判定规则,将元器件轮廓检测扩展到抓取点和装配点定位,得到丰富的类别和位置指引信息,为工业机器人精准、可靠、稳定地抓取和装配做好基础。

下一步研究准备进行相机和机械臂的标定,加入深度



矩阵得到各元器件深度值;构建强耦合机器人视觉控制模型;设计工业机器人视觉导向的变刚度柔顺控制方法,进而实现元器件智能装配机器人的智能装配。

### 参考文献

- [1] 张志杰,顾寄南,李静,等. 基于深度学习的电子元器件快速检测算法研究[J]. 电子测量技术, 2022, 45(10):93-101.
- [2] 周济. 智能制造——“中国制造 2025”的主攻方向[J]. 中国机械工程, 2015, 26(17):2273-2284.
- [3] BHAGYA C, SHYNA A. An overview of deep learning based object detection techniques[C]. 2019 1st International Conference on Innovations in Information and Communication Technology, Chennai, 2019: 1-6.
- [4] 侯学良,单腾飞,薛靖国. 深度学习的目标检测典型算法及其应用现状分析[J]. 国外电子测量技术, 2022, 41(6):165-174.
- [5] ZAIDI S S A, ANSAEI M S, ASLAM A, et al. A survey of modern deep learning based object detection models[J]. Digital Signal Processing, 2022: 103514.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision. Springer, Cham, 2016: 21-37.
- [7] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [8] 周志华. 机器学习[M]. 北京:清华大学出版社, 2016.
- [9] YANG X, YAN J, FENG Z, et al. R3det: Refined single-stage detector with feature refinement for rotating object [C]. Proceedings of the AAAI Conference on Artificial Intelligence, California, 2021, 35(4): 3163-3171.
- [10] HOU Y, SHI G, ZHAO Y, et al. R-YOLO: A YOLO-based method for arbitrary-oriented target detection in high-resolution remote sensing images[J]. Sensors, 2022, 22(15): 5716.
- [11] LIU Z, HU J, WENG L, et al. Rotated region based CNN for ship detection[C]. 2017 IEEE International Conference on Image Processing, Beijing, 2017: 900-904.
- [12] QIAN W, YANG X, PENG S, et al. Learning modulated loss for rotated object detection [C]. Proceedings of the AAAI Conference on Artificial Intelligence, California, 2021, 35(3): 2458-2466.
- [13] 李壮飞,杨风暴,郝岳强. 一种基于残差网络优化的航拍小目标检测算法[J]. 国外电子测量技术, 2022, 41(8):27-33.
- [14] 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, Las Vegas, 2016: 770-778.
- [15] YU F, WANG D, SHELHAMER E, et al. Deep layer aggregation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, 2018: 2403-2412.
- [16] 杨瑾,陈灯,张彦铎,等. 基于多尺度特征融合网络的铁路工机具目标检测[J]. 电子测量技术, 2022, 45(17): 94-100.
- [17] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More Features from cheap operations [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern, Recognition, 2020: 1580-1589.
- [18] 刘涛,张涛. 基于 GhostNet-YOLOv4 算法的印刷电路板缺陷检测[J]. 电子测量技术, 2022, 45(16):61-70.
- [19] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 764-773.

### 作者简介

雷文桐,硕士研究生,主要研究方向为机器视觉、深度学习、人工智能等。

E-mail:1091653438@qq.com

顾寄南(通信作者),博士,教授,博士生导师,主要研究方向为智能制造系统、智能机器人、机器视觉等。

E-mail:gjn@mail. ujs. edu. cn

胡君杰,硕士研究生,主要研究方向为深度学习、目标检测。

高伟,硕士研究生,主要研究方向为基于机器视觉的缺陷检测。