

# 基于改进 FoveaBox 的废杂塑料检测

文生平<sup>1,2</sup>, 陈敬福<sup>1,2</sup>, 冯泽锋<sup>1,2</sup>, 朱珂郁<sup>3</sup>

(1. 华南理工大学 广东省高分子先进制造技术及装备重点实验室, 广州 510640;

2. 华南理工大学 聚合物成型加工工程教育部重点实验室, 广州 510640;

3. 福建师范大学 环境科学与工程学院, 福州 350007)

**摘要:** 塑料制品的产量和种类的飞速增长, 给废杂塑料的回收带来极大的挑战; 目前面对恶劣和高强度的工作环境, 仍然依靠大量人工分拣, 无疑亟待自动化升级; 为解决上述问题, 提出了一种改进的 FoveaBox 目标检测算法; 针对废杂塑料分选背景复杂的问题, 采用 ResNeXt-101 作为主干网络替代 ResNet-50 来提高特征提取能力; 针对外形差异大的问题, 采用带缩放系数的可变形卷积来提高卷积过程的有效感受野; 针对目标间彼此遮挡问题, 采用带层级控制因子的软化加权锚点机制来提高被遮挡目标的检测精度; 结果表明, 基于改进 FoveaBox 的废杂塑料检测算法检测平均精度均值达到 85.79%, 检测速度为 71.4 ms, 该检测算法较强的实用性得到验证。

**关键词:** 塑料分选; 目标检测; 改进 FoveaBox; 可变形卷积; 软化加权锚点

## Detection of Waste Plastic Based on improved FoveaBox

WEN Shengping<sup>1,2</sup>, CHEN Jingfu<sup>1,2</sup>, FENG Zefeng<sup>1,2</sup>, ZHU Keyu<sup>3</sup>

(1. Guangdong Provincial Key Laboratory of Technique and Equipment for Macromolecular Advanced Manufacturing, South China University of Technology, Guangzhou 510640, China;

2. Ministerial Key Laboratory of Polymer Processing Engineering, South China University of Technology, Guangzhou 510640, China;

3. College of Environmental Science and Engineering, Fujian Normal University, Fuzhou 350007, China)

**Abstract:** The rapid growth in the output and types of plastic products has brought great challenges to the recycling of waste and miscellaneous plastics. At present, facing the harsh and high-intensity working environment, it still relies on mass manual sorting, it is undoubtedly urgent to upgrade automation. In order to solve above problems, an improved FoveaBox target detection algorithm is proposed. In view of the complicated background for waste plastic sorting, ResNeXt-101 is used as the backbone network to replace ResNet-50 to improve the feature extraction ability. Aiming at the problem of large shape differences, a deformable convolution with a zoom factor is used to improve the effective receptive field in the processing of convolution. Aiming at the mutual occlusion problem between targets, a softening weighted anchor point mechanism with hierarchical control factors is used to improve the detect accuracy of obscured targets. The results show that the mean average precision of the waste plastic detection algorithm based on the improved FoveaBox reaches by 85.79%, and the detection speed is up to 71.4 ms, the detection algorithm is strong practically verified.

**Keywords:** plastic classification; target detection; advanced FoveaBox; deformable convolution; soft-weighted anchor point

## 0 引言

塑料制品在生产与生活的各个领域都有广泛的应用。随着经济的发展和人们需求的多样化, 塑料制品的种类与产量飞速上升, 同时也给废杂塑料的回收带来巨大的挑战。废杂塑料回收利用的第一步是从众多可回收垃圾中按照形状、颜色、用途等对整件废杂塑料制品进行分类, 以方便后续压扁运输以及破碎精分等回收操作<sup>[1]</sup>。当前我国对废杂塑料进行整件分选的产业线仍处于人工作业的低效率模式, 恶劣的工作环境、高昂的人工费用、低效的分选速度

以及不稳的分选精度严重制约着产业的发展, 废杂塑料分拣行业无疑亟需从人工到自动化的转型<sup>[2]</sup>。然而, 废杂塑料因分选对象繁多、外形尺寸差异大、存在彼此遮挡、污染破损等因素影响, 基于传统机器视觉的自动分选方法难以有效应对如此复杂的场景。

人工智能的发展为机器视觉领域带来新思路, 已有不少研究人员将深度学习应用于各种消费与工业场景。肖体刚等人以 Yolov3 模型为基础, 融合深度可分离卷积结构和多尺度特征检测, 实现对安全帽的智能监控检测<sup>[3]</sup>; 李震霄等人提出 Yolov3-MobileNetv2 算法, 解决多车辆跟踪场景中的实时

收稿日期: 2021-10-15; 修回日期: 2021-11-19。

基金项目: 国家重点研发计划(2019YFC1908201)。

作者简介: 文生平(1966-), 男, 湖北荆州人, 工学博士, 教授, 主要从事工业装备的智能控制与机器视觉方向的研究。

引用格式: 文生平, 陈敬福, 冯泽锋, 等. 基于改进 FoveaBox 的废杂塑料检测[J]. 计算机测量与控制, 2022, 30(4): 66-71.

性和身份跳变问题<sup>[4]</sup>; 董旭彬利用 SD-RPN 改进 Mask R-CNN 算法, 且融合可见光图像与红外图像以解决航空影像目标检测中的性能缺陷问题<sup>[5]</sup>; 牛作东引入自注意力机制改进 RetinaFace 算法, 提出了自然场景下人脸口罩佩戴与否的检测思路<sup>[6]</sup>; 文生平等人基于 Gaussian-Yolov3 算法, 实现了对铝型材表面缺陷的实时检测<sup>[7]</sup>。

目前针对废杂塑料检测算法的研究仍较少。赵远远等人提出基于 Zernike 矩的图像轮廓特征提取及欧氏距离分类算法, 解决重叠塑料瓶对废旧塑料瓶颜色分选的干扰<sup>[8]</sup>; 阎嘉旭等人利用 Yolov3 模型对绿、蓝、白和透明四色废旧塑料瓶进行检测<sup>[9]</sup>; 张海洲等人同样对不同颜色的塑料瓶进行分选, 但对 Yolov3 模型的主干网络进行修改, 提出了 Bottle-Yolov3 算法<sup>[10]</sup>。目前已有研究大多按颜色对废旧塑料瓶进行分选, 局限于塑料瓶这一小对象, 场景比较理想化, 且以颜色进行分类难以满足当下生产需求。

然而, 采用普通的目标检测模型难以有效应对废杂塑料分拣场景, 第一个原因是废杂塑料分选背景复杂, 实际分选流水线往往由于粗筛不彻底, 有其他生活垃圾混入干扰, 同时生活中废杂塑料产品种类繁多, 且后期使用中还存在被污染、破损、碾压, 物品外观已发生较大变化, 因此需要模型具有出色的特征提取能力; 第二个原因是不同物品间外形差异较大, 比如在尺度大小和宽高比例方面存在较大差距。第三个原因是尽管预处理过程会尽量将物品振荡平铺, 但还是无法避免物品间彼此遮挡, 模型可能会漏检被遮挡的物体。为此, 本文将针对饮料瓶、洗漱用品瓶、一次性餐具盒 3 个更大粒度的废杂塑料对象进行检测, 选取 FoveaBox<sup>[11]</sup>作为目标检测的基础模型。为了提高模型在复杂分选背景下的特征提取能力, 采用更深、更宽的 ResNeXt-101<sup>[12]</sup>作为主干网络代替 ResNet-50。为了解决废杂塑料外形差异大的问题, 采用带有缩放系数的可变形卷积代替普通卷积以提升卷积过程的有效感受野。为了解决物品遮挡的问题, 采用带层级控制因子的软化加权锚点机制, 减少被遮挡目标的漏检率。经过以上改进, 使得模型更适用于废杂塑料分选场景, 为后续实验线的运行提供了算法支持。

### 1 基于经典 FoveaBox 的废杂塑料检测流程

废杂塑料检测需要一种能满足复杂场景下多目标的实时检测算法。FoveaBox 模型通过主干网络进行特征提取, 再由特征金字塔对不同尺寸的特征图像进行上采样和特征融合, 增强了网络对复杂场景下废杂塑料的检测能力。在多次特征融合后, 会得到不同大小的检测头, 实现对大小不一的塑料对象的有效检测。

不同于 Faster-RCNN<sup>[13]</sup>、TridentNet<sup>[14]</sup> 等二阶段 (two-stage) 检测器, FoveaBox 无需经过候选区域, 通过采用一阶段 (one-stage) 路线, 利用检测头直接求得物体的类别概率和位置坐标值; 同时不像 Yolo、SSD<sup>[15]</sup>、RetinaNet<sup>[16]</sup> 等基于锚框 (anchor-based) 的模型, FoveaBox 采

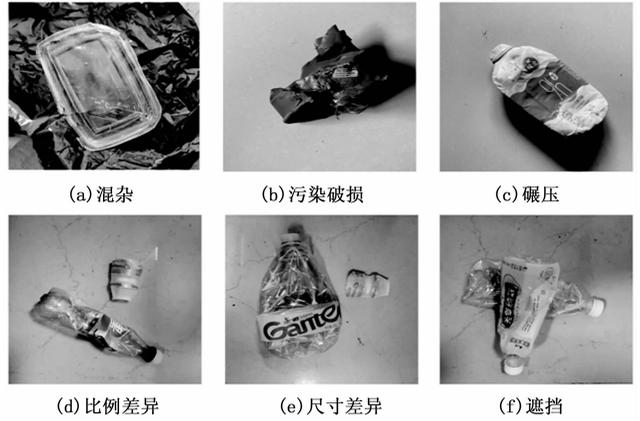


图 1 各类分选场景

用无锚框 (anchor-free) 的思想, 摒弃了大量密集的先验框。而无锚框的检测模型又可粗略分为两类——基于锚点 (anchor-point) 检测和关键点 (key-point) 检测。基于锚点检测的模型有 FSAP<sup>[17]</sup>, FCOS<sup>[18]</sup> 和 FoveaBox 等, 其将检测框编码成锚点; 而基于关键点检测的模型有 CornerNet<sup>[19]</sup>, ExtermeNet<sup>[20]</sup> 和 CenterNet<sup>[21]</sup> 等, 其将检测问题重新定义为检测框的一些关键点定位问题, 如角点、中心点或极点。基于关键点检测模型依赖于单个高分辨率特征图和重复的底到顶、顶到底推理, 因此基于锚点检测的模型往往有更短的训练和推理时间。FoveaBox 作为一种基于锚点的一阶段无锚框的检测模型, 减少了网络模型中的大量参数, 保证废杂塑料检测的实时性。

FoveaBox 与其它基于锚点检测模型最大区别在于预测目标的区域只取目标中心区域 (即中心凹, fovea) 的锚点, 而非目标所在区域的所有锚点, 中心凹大小为原区域的 0.4。网络结构如图 2 所示, 其中主干网络为 ResNet-50, 经过每层特征金字塔  $P_l$  融合后各会输出一个检测头, 检测头特征图尺寸为  $W \times H/s_l$ ,  $s_l$  为下采样率,  $s_l = 2^l$ ,  $l$  的取值为 3 ~ 7, 即一共有 5 个不同尺寸的检测头。每个检测头包含分类子网络和边框回归子网络, 其中分类子网络会对每个锚点预测属于  $K$  个类别的概率, 边框回归子网络则为每个锚点预测与类别无关的目标位置。

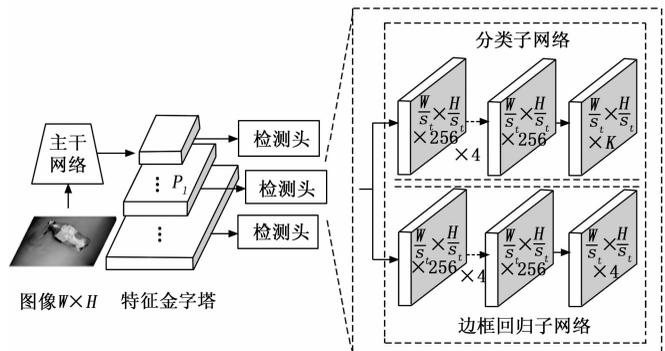


图 2 FoveaBox 模型结构

### 2 改进方案——可变形卷积

卷积神经网络对于未知形状变换建模的固有缺陷来源于其基础模块固有的几何结构，这种模式下卷积过程中每一个像素点的感受野是固定的，无法针对待检测目标的特殊性自适应的改变采样的结构。在废杂塑料制品的整件视觉分选过程中，待分选目标的外形尺寸存在较大差异，有些方正有些狭长，从而降低模型的检测精度。

为解决这种建模上的局限性，将 FoveaBox 中的普通卷积替换为可变形卷积<sup>[22]</sup>。通过添加每个采样点的偏移量，利用反向传播算法的迭代优化学习偏移量，以使卷积的感受野发生改变，获取更灵活的采样区域。

传统卷积中，以  $3 \times 3$  卷积核为例，对于某一卷积输出  $p_0$ ，都是从以  $x(p_0)$  为中心扩散的 9 个像素点区域  $R$  采样，进而与卷积核  $\omega$  做乘积和得出的结果。采样区域以中心采样点偏置的表示方式如式 (1) 所示：

$$R = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\} \quad (1)$$

传统卷积的输出公式如下所示：

$$y(p_0) = \sum_{p_n \in R} \omega(p_n) \cdot x(p_0 + p_n) \quad (2)$$

可变形卷的输出公式如下所示：

$$y(p_0) = \sum_{p_n \in R} \omega(p_n) \cdot x(p_0 + p_n + \Delta p_n) \quad (3)$$

其中： $\Delta p_n$  为每个采样位点的偏移量，偏移量是在原有卷积的基础上增添一条卷积路径以学习其中参数的。为避免偏移量过大学习大量无用的上下文信息，反向干扰特征提取信息的有效性，因此对可变形卷积进行优化，限制偏移量的幅值。优化后的可变形卷积如式 (4) ~ (5) 所示：

$$\eta = (\alpha(p_n^x - p_0^x), \beta(p_n^y - p_0^y)) \quad (4)$$

$$y(p_0) = \sum_{p_n \in R} \omega(p_n) \cdot x(p_0 + p_n + \Delta p_n \cdot \eta) \quad (5)$$

其中： $\eta$  为缩放系数， $\alpha$  和  $\beta$  为待学习超参数，初始化为 2。

可变形卷积结构如图 3 所示，其中，增添一条卷积路径，以学习缩放系数，缩放系数的效果是根据原采样位置，对变形后的采样位置进行调整，使得最终的采样结果在像素层面上保持连贯性，在语义层面上保持依赖性。

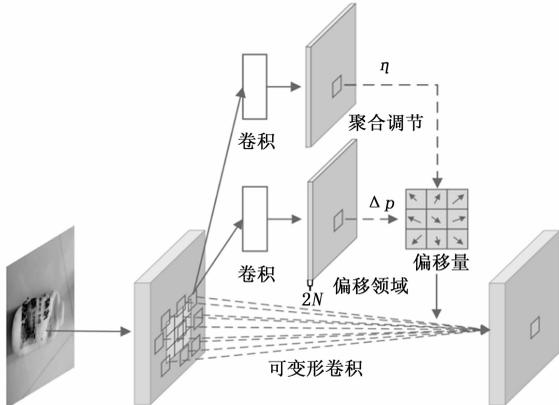


图 3 可变形卷积结构图

### 3 改进方案——软化加权锚点机制

基于锚点的检测模型 FoveaBox 面对物体遮挡的复杂场景，会因为特征不对齐而导致目标边界位置得到不必要的高分，抑制被遮挡物体的评分，从而使得模型将注意力偏向于靠前更完整的物体，忽略了被遮挡物体的检测，即注意力偏差问题。

图 4 为输入图像和分类子网络输出的得分热力图。图中存在两个待检测物品，靠前完整的一次性包装盒所在区域生成一个很大高分区域，并且倾向于向被遮挡的饮料瓶区域扩展，严重情况下，会造成饮料瓶的漏检。虽然废杂塑料分拣流水线会使用振荡等预处理方式让物品平铺，彼此尽可能分散，但难免会出现塑料目标存在紧挨、互相遮挡的情况。为了解决注意力偏差问题，引入软化加权锚点机制<sup>[23]</sup>，根据锚点到目标中心的距离赋予不同的权重，避免物体边缘和中心的锚点被同等对待。

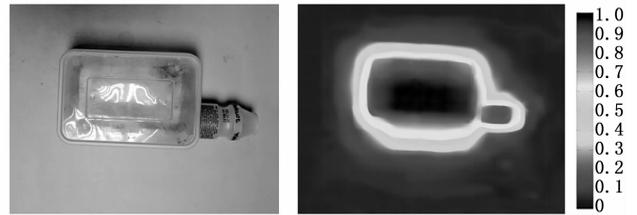


图 4 输入图像与特征热力图

在 FoveaBox 模型中，锚点  $p_{i,j}$  是金字塔特征图  $P_l$  上  $(i, j)$  位置的像素点，每个锚点在原图上对应一个图像空间位置  $(X_{i,j}, Y_{i,j})$ ，相应计算公式如式 (6) ~ (7) 所示：

$$X_{i,j} = s_l(i + 0.5) \quad (6)$$

$$Y_{i,j} = s_l(j + 0.5) \quad (7)$$

当锚点在标签框被分配到的特征金字塔  $P_l$  层时且  $(X_{i,j}, Y_{i,j})$  在目标中央凹内时，该锚点才为前景锚点，否则都是背景锚点。

前景锚点为对于原图上某个目标实例的标注框在金字塔特征图上的某一区域，如果锚点落入当前该区域以内则判定为前景锚点，否则为后景锚点。前景锚点的位置编码标签为  $(d^l, d^t, d^r, d^b)$ ，分别对应锚点相对于标签框的左、上、右、下。其计算公式如式 (8) ~ (11) 所示：

$$d^l = \frac{1}{zs_l} [X_{i,j} - (x - \frac{w}{2})] \quad (8)$$

$$d^r = \frac{1}{zs_l} [(x + \frac{w}{2}) - X_{i,j}] \quad (9)$$

$$d^t = \frac{1}{zs_l} [Y_{i,j} - (y - \frac{h}{2})] \quad (10)$$

$$d^b = \frac{1}{zs_l} [(y + \frac{h}{2}) - Y_{i,j}] \quad (11)$$

式中， $z$  为标准化因子， $(x, y, w, h)$  为目标实例在原图上的标签框坐标信息。

实现权重软化的方案是多样的，最简单的为采用锚点到目标中心距离单调递减的函数，本文则采用广义中心度

函数<sup>[17]</sup>, 公式如下所示:

$$\omega_{l_{ij}} = \begin{cases} \left[ \frac{\min(d_{l_{ij}}^l, d_{l_{ij}}^r) \cdot \min(d_{l_{ij}}^l, d_{l_{ij}}^b)}{\max(d_{l_{ij}}^l, d_{l_{ij}}^r) \cdot \max(d_{l_{ij}}^l, d_{l_{ij}}^b)} \right]^{\eta^L}, p_{l_{ij}} \in p^+ \\ 1, p_{l_{ij}} \in p^- \end{cases} \quad (12)$$

其中:  $\eta$  为缩放比率, 控制权重递减的幅度,  $(L - l)/L$  为层级控制因子, 控制不同层次的特征金字塔的权重进行差别性衰减,  $p^+$  和  $p^-$  分别为前景与背景锚点的集合, 权重  $\omega_{l_{ij}}$  的取值范围为  $0 \sim 1$ 。

软化加权锚点机制如图 5 所示, 对于同层的前景锚点, 锚点越接近预测框的中心, 被赋予更大的权重, 锚点越接近预测框边缘, 则被赋予较小的权重以减轻对该锚点的关注度; 对于不同层级的前景锚点, 由于深层的特征图主要检测大尺寸目标, 可能会跟其他目标具有较大面积的重叠情况, 因此需要加速其权重衰减的速率, 而浅层的特征图主要检测小尺寸目标, 如果设置过高的衰减比率, 则会导致前景锚点数量不足, 使得模型受到较大的正负样本不平衡的压力, 因此需要减缓其权重衰减的速率, 层级控制因子则很好地解决了这个问题。

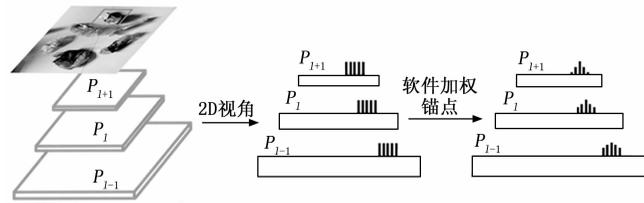


图 5 软化加权锚点机制示意图

## 4 实验结果及分析

### 4.1 数据集与实验软硬件配置

本数据集主要分为 3 个类别, 包括饮料瓶、洗漱用品瓶和一次性餐具盒, 以实地采集与网络爬虫的形式相结合, 其中通过相关图片网站爬取筛选 6 000 张图片, 同时实地采集 3 000 张待分选废杂塑料图片, 部分图片样例如图 6 所示。针对算法训练的要求, 将打乱后的数据集以 7 : 1.5 : 1.5 的比例随机划分为训练数据、验证数据以及测试数据, 且保证各个数据集中各类别的比例相近。为了更好地训练网络参数, 防止模型过拟合, 利用数据增强的方式对已有数据集进行扩充, 如旋转、翻转、裁剪、模糊、添加噪声等。

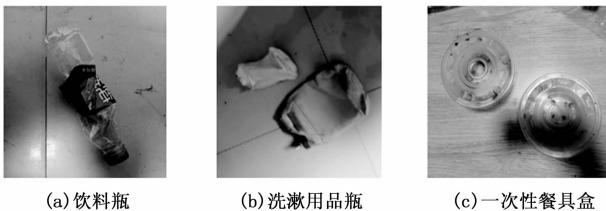


图 6 废杂塑料图像数据示例

实验主要软硬件配置如表 1 所示。

表 1 软硬件配置表

硬件	CPU	Intel(R) Core(TM) i7-7700
	GPU	NVIDIA GeForce GTX 3070
	内存	32 GB
软件	开发语言	Python3.7
	操作系统	Ubuntu18.04
	深度学习框架	Pytorch
	图像标注工具	LabelImg

### 4.2 评价指标

$P-R$  曲线是以召回率为横坐标, 精确率为纵坐标构成的曲线图像, 如图 7 所示。其是将模型对于分类的阈值从低往高逐次提高过程中  $P$  与  $R$  的值作为坐标点连接而成的曲线, 直观表示了分类器的性能, 一般认为  $P-R$  曲线与坐标轴包围的面积越大, 模型性能越佳。其中平衡点表示  $P=R$  时的情况。

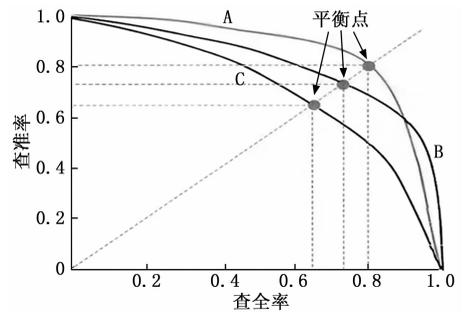


图 7  $P-R$  曲线示意图

在本文废杂塑料检测任务中, 将使用 PASCAL VOC 平均精度均值 (mAP, mean average precision) 作为模型算法的检测性能指标, mAP 衡量的是模型在所有类别上分类性能的总体评价。同时, 使用模型对一张废杂塑料图像的检测时间作为算法的检测效率指标。如公式 (13) 所示:

$$mAP = \frac{\sum_{i=1}^Q (AP)_i}{Q} \quad (13)$$

式中,  $Q$  表示类别数量,  $AP$  为平均精度, 表示  $P-R$  曲线与  $x, y$  轴的封闭区域面积, 表示模型在一个类别上的分类性能。

### 4.3 主干网络对比实验

在经典的 FoveaBox 中, 主干网络为 ResNet-50, 其中的残差模块能有效抑制梯度消失以及爆炸的现象, 在深度学习领域具有良好的应用范围。ResNeXt 残差模块对同一特征图采用不同大小的卷积核做卷积操作再拼接以融合不同尺度的特征, 同时以多个  $3 \times 3$  小卷积叠加代替大卷积的方法, 使得 ResNeXt 残差模块与 ResNet 残差模块在同等网络深度下, 具有更宽的通道, 但参数量却相差不大, 因而具有更好的表现。

ResNeXt 模型与 ResNet 模型是以同一拓扑结构聚合的方式构建卷积模块, 后面的数字代表不同的卷积层数。利用同

样的训练与测试图像数据，通过对改进 FoveaBox 替换不同的特征网络进行性能对比，以探索何种方案对本文废杂塑料目标检测场景更具适用性。

表 2 不同主干网络对改进 FoveaBox 算法的效果对比

主干网络	$mAP/\%$	检测时间/ms
ResNet-50	78.24	54.4
ResNet-101	79.78	68.8
ResNeXt-50	79.33	57.5
ResNeXt-101	81.75	71.4

从表 2 的数据可以看出，增加网络深度能有效提高  $mAP$ ，但也需花费更长的检测时间；采用 ResNeXt 残差模块的模型在同等网络深度下，只需增加少量检测时间的代价，模型就有更好的表现，证明了增加网络的宽度也是提高  $mAP$  的有效途径。从对比实验也可以看出，ResNeXt-101 在增加 17 ms 的检测时间情况下， $mAP$  比 ResNet-50 多了 3.51%，在保证检测实时性的基础上具有更强的特征提取能力，能更有效应对复杂的废杂塑料分拣场景。

#### 4.4 数据增强性能实验

网络爬取的大量相关图像并不完全匹配废杂塑料视觉分选场景，同时，产品种类繁多及使用运输过程中造成的外观变化，需要采集更多的图像以覆盖各种场景。一般而言，训练的图像数据量越大，模型训练效果越好，因此有必要对图像集进行扩充。数据增强是视觉领域提升网络训练效果最常见且有效的途径。常用的数据增强包括旋转、水平翻转、裁剪、噪声、模糊等。

本文将设计多种数据增强的方案，通过对已有图像数据的多种变换，扩充训练数据的规模，以探索在本文废杂塑料检测的中小型数据集下，引入人工数据增强的方案。下面通过实验对比对不同的数据增强方案加以验证其实际效果，实验效果如表 3 所示。其中，对于每一张图像，设置其将有 30% 的概率分别进行 10°、15°、20°、25°、30° 的旋转，40% 的概率进行水平翻转，10% 的概率进行裁剪，20% 的概率添加噪声，10% 的概率进行模糊。

表 3 数据增强实验对比

数据增强策略	$mAP/\%$
无数据增强	81.75
旋转	82.45
旋转+水平翻转	83.16
旋转+水平翻转+裁剪	83.53
旋转+水平翻转+裁剪+噪声	84.31
旋转+水平翻转+裁剪+噪声+模糊	85.79

#### 4.5 改进方案消融实验

为了直观感受使用可变形卷积的效果变化，对卷积过程中采样位置进行记录并显示，从而比较模型在采样过程中的感受野变化。传统卷积与可变形卷积的采样位置如下图 8 所示。先在顶层的特征图选取一个激活点在饮料瓶处，

经过多次卷积后，传统卷积的采样位置仅局限于框定的正方形中，无法较好的适应待检测目标的形态与位姿，而应用可变形卷积后，经过多层的卷积后，深层特征图的一个采样点将扩散至饮料瓶的不同区域，同时由于限制偏移量的幅值，扩散范围也有所限制，从而能针对性的对待检测目标进行特征提取。

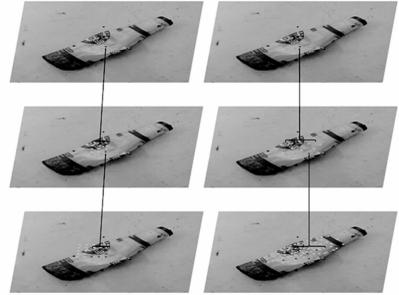


图 8 传统卷积与可变形卷积采样过程

同样，对使用软化加权锚点机制的特征热力图进行对比，如图 9 所示。与图 4 相比，使用软化加权锚点机制后，一次性包装盒所在的高分区域明显变小，边缘锚点权重接近于零，两个目标都生成了较为明显的得分区域，两者之间也能有效地进行区分，注意力偏差问题得到较好的解决。

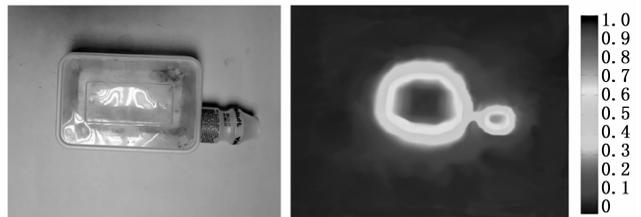


图 9 输入图像与软化加权锚点后的特征热力图

为进一步验证上文提及各改进方案对废杂塑料分选场景的有效性，下面将通过消融实验加以验证，以  $mAP$  和检测时间作为评判依据。

表 4 消融实验

改进方案	可变形卷积	软化加权锚点	$mAP/\%$	检测时间/ms
ResNeXt-101 + FoveaBox	×	×	80.84	68.3
	√	×	82.96	70.8
	×	√	83.52	69.2
	√	√	85.79	71.4

如表 4 所示，在原始 FoveaBox 算法中， $mAP$  值仅有 80.84%，但检测效率达到 68.3 ms。两种改进方案均能提高检测效果，但由于增加了计算量，均不同程度增加了检测时间。其中，使用软化加权锚点机制对于模型的精度提升最高，但少量增加了检测时间。

对于融合两种改进方案的 FoveaBox 模型，可以发现改进方案之间并没有明显的效果冲突现象，最后  $mAP$  达到 85.79%，平均检测效率实现 71.4 ms 检测一张图像的效果。

相对于其他方案都具有一定的提升,且满足工业场景实时作业的效率要求,验证了本文提出的改进 FoveaBox 算法的有效性与其可行性。

## 5 结束语

本文针对废杂塑料的分选场景,提出了一种基于改进的 FoveaBox 目标检测算法。为了解决复杂背景下目标检测问题,使用 ResNeXt-101 作为主干网络以提高特征提取能力。为了解决废杂塑料外形差异大的问题,使用带缩放系数的可变形卷积来提高卷积过程的有效感受野。为了解决目标间彼此遮挡的问题,使用带层级控制因子的软化加权锚点机制,使模型更有效地区分紧挨的物品。通过消融实验,证明改进方案的有效性,改进模型在满足实时性的情况,具有更高的检测精度,基本满足实际生产需求。随着采集的样本数据日渐丰富,模型检测能力将进一步增强。

## 参考文献:

- [1] 王海南,傅和清. 废塑料的回收利用技术进展 [J]. 黄冈职业技术学院学报, 1999 (4): 65-69.
- [2] 宋林,张杨. 创新驱动下制造业的产业转型升级 [J]. 西安交通大学学报(社会科学版), 2020, 40 (1): 38-47.
- [3] 钟鑫豪,龙永红,何震凯,等. 基于改进 Tiny-yolov3 算法的安全帽佩戴检测 [J]. 湖南工业大学学报, 2021, 35 (2): 46-50.
- [4] 李震霄,孙伟,刘明明,等. 交通监控场景中的车辆检测与跟踪算法研究 [J]. 计算机工程与应用, 2021, 57 (8): 103-111.
- [5] 董旭彬,赵清华. 改进 Mask R-CNN 在航空影像目标检测的研究应用 [J]. 计算机工程与应用, 2021, 57 (8): 133-144.
- [6] 牛作东,覃涛,李捍东,等. 改进 RetinaFace 的自然场景口罩佩戴检测算法 [J]. 计算机工程与应用, 2020, 56 (12): 1-7.
- [7] 文生平,李超贤. 基于 Gaussian-yolov3 的铝型材表面缺陷检测 [J]. 计算机测量与控制, 2020, 28 (9): 88-93.
- [8] 赵远远. 基于计算机视觉的废旧塑料瓶分拣回收系统 [D]. 杭州: 电子科技大学, 2020.
- [9] 阎嘉旭. 废旧塑料瓶智能分拣技术研究 [D]. 杭州: 电子科技大学, 2020.
- [10] 张海洲. 基于深度学习的废旧塑料饮料瓶分选系统研究 [D]. 南京: 南京理工大学, 2020.

(上接第 65 页)

- [15] GUI X S, HONG J T, DAI H J, et al. A signal modulation recognition method based on convolutional neural network [J]. Modern Computer, 2019 (10): 18-22, 26.
- [16] 王锋涛,敖银辉. 基于多重分形理论的地铁车辆平轮故障诊断 [J]. 机械工程与自动化, 2018 (5): 12-14.
- [17] 李红延,周云龙,田峰,等. 一种新的小波自适应阈值函数振动信号去噪算法 [J]. 仪器仪表学报, 2015, 36 (10): 2200-2206.

- [11] KONG T, SUN F, LIU H, et al. FoveaBox: beyond anchor-based object detector [J]. arXiv preprint arXiv: 1904.03797, 2019.
- [12] XIE S, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks [C] // IEEE. IEEE, 2016.
- [13] GaVRILESCU R, ZET C, FOALU C, et al. Faster R-CNN: an approach to real-time object detection [C] // 2018 International Conference and Exposition on Electrical and Power Engineering (EPE). IEEE, 2018.
- [14] LI Y, CHEN Y, WANG N, et al. Scale-aware trident networks for object detection [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019.
- [15] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [J]. arXiv preprint arXiv: 1512.02325, 2016.
- [16] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017 (99): 2999-3007.
- [17] ZHU C, HE Y, SAVVIDES M. Feature selective anchor-free module for single-shot object detection [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019.
- [18] TIAN Z, SHEN C, CHEN H, et al. FCOS: fully convolutional one-stage object detection [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, 2019.
- [19] LAW H, DENG J. CornerNet: detecting objects as paired keypoints [J]. International Journal of Computer Vision, 2020, 128 (3): 642-656.
- [20] ZHOU X, ZHUO J, PHILIPP K. Bottom-up object detection by grouping extreme and center points [J]. arXiv preprint arXiv: 1901.08043, 2019.
- [21] DUAN K, BAI S, XIE L, et al. CenterNet: Object Detection with Keypoint Triplets [J]. arXiv preprint arXiv: 1904.08189, 2019.
- [22] DAI J, QI H, XIONG Y, et al. Deformable convolutional networks [J]. arXiv Preprint arXiv: 1703.06211, 2017.
- [23] ZHU C, CHEN F, SHEN Z, et al. Soft anchor-point object detection [M]. Springer, Cham, 2020.

- [18] HE L F, CUI Y Y, ZHANG T Q, et al. Analysis of weak signal detection based on tri-stable system under Levy noise [J]. Chinese Physics B, 2016, 25 (6): 85-94.
- [19] 张刚,杨玉蕾,张天骥. Levy 噪声下自适应级联三稳随机共振系统研究 [J]. 电子测量与仪器学报, 2019, 33 (4): 160-168.
- [20] WERON R. On the Chambers-Mallows-Stuck method for simulating skewed stable random variables [J]. Statistics & Probability Letters, 1996, 28 (2): 165-171.