

基于 Mask RCNN 的滤袋开口检测方法

王宪保, 朱啸咏, 姚明海

(浙江工业大学 信息工程学院, 杭州 310023)

摘要: 滤袋开口的检测与定位在滤袋智能生产过程中占据着至关重要的地位; 但由于滤袋具有柔性的特点, 常规检测方法很难有效进行, 且定位精度也不能满足生产要求; 针对此问题, 提出一种结合可变形卷积与掩码信息的多尺度目标检测器, 该检测器使用可变形卷积改进主干网络高层中的固定卷积, 结合特征金字塔技术实现多尺度信息融合; 然后将所得多尺度信息通过区域提议网络生成候选区域, 采用改进的 Soft-NMS 方法进行筛选, 最终送入检测头进行识别与分割; 通过滤袋图像数据集进行实验; 结果表明, 提出的算法相较于基准算法有了 2.4 个百分点的检测精度提升, 实现了滤袋开口的准确识别与高精度定位。

关键词: 目标检测; 可变形卷积; Mask RCNN; 改进 Soft-NMS

Detection Method of Filter Bag Opening Based on Mask RCNN

Wang Xianbao, Zhu Xiaoyong, Yao Minghai

(Department of Information, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: The detection and positioning of the opening of the filter bag occupies a crucial position in the intelligent production process of the filter bag. However, due to the flexible characteristics of the filter bag, the conventional detection method is difficult to carry out effectively, and the positioning accuracy cannot meet the production requirements. To solve this problem, a multi-scale target detector combining deformable convolution and mask information is proposed. The detector uses deformable convolution to improve the fixed convolution in the high layer of the backbone network, combined with feature pyramid technology to achieve multi-scale information fusion. Then, the obtained multi-scale information is generated through the region proposal network to generate candidate regions, which are screened using the improved Soft-NMS method, and finally sent to the detection head for recognition and segmentation. The experiment is carried out through the filter bag image data set. The results show that the proposed algorithm has a detection accuracy improvement of 2.4% compared to the benchmark algorithm, and achieves accurate identification and high-precision positioning of the filter bag opening.

Keywords: target detection; deformable convolutional networks; mask RCNN; improved Soft-NMS

0 引言

在滤袋的智能生产中, 需要对滤袋进行卡扣的自动化安装, 而实现这一过程的前提是获得滤袋开口的准确位置及滤袋所属的类别。伴随着滤袋制造工艺的提升以及需求的多样化, 滤袋大小、纹理等变得更加丰富, 这为滤袋开口检测问题带来了更为严峻的挑战。

本文的滤袋开口检测实际上是一种目标检测, 它是一种包含目标定位与目标分类任务, 使用图像处理、机器学习等方法, 在图像中搜索感兴趣对象的计算机视觉技术^[1]。一般先通过目标分类判断是否含有目标对象, 再进行更细粒度的检测, 并将搜索到的目标通过方框进行标注^[2]。

计算机视觉处理中, 图像一般以像素矩阵的形式存储,

对其进行类别以及位置信息的相关特征提取方可实现目标检测^[3]。传统目标检测算法在图像经过去噪、增强、裁剪等预处理之后, 对图像采用滑动窗口遍历并生成候选区域, 对每个候选区域采取特征提取, 例如方向梯度直方图 (HOG, histogram of oriented gradients)^[4], 尺度不变特征变换 (SIFT, scale invariant feature transform)^[5], (DPM, deformable parts model)^[6] 等, 最后通过 AdaBoost^[7], (SVM, support vector machine)^[8] 等机器学习算法对目标进行分类, 然后依据类别对目标进行边框回归。由于上述目标检测模型对于不同的特征需要设计不同分类器, 导致其泛用性差、鲁棒性不足。

2012 年之后, 计算机视觉技术在深度学习的带动下开始了迅猛的发展, 以深度卷积神经网络为代表的方法替代了手工特征方法。在目标检测领域, (Fast RCNN, fast region-based convolutional neural network)^[9] 和 (Faster RCNN, faster region-based convolutional neural network)^[10] 为基于深度学习的目标检测方法提供了新的方向。前者由于共享卷积操作降低了网络整体的计算消耗, 后者在前者基础上融合全卷积网络^[11] 的思想, 提出通过区域提议网络 (RPN, region proposal networks) 代替原先的 Selective Search^[12] 以及 Edge Boxes^[13] 算法。另一种基于回归的思路由 2016 年 Redmon 等^[14] 提出的

收稿日期: 2020-04-30; 修回日期: 2020-05-13。

基金项目: 国家自然科学基金项目 (61871350); 浙江省科技计划项目 (2019C011123); 浙江省基础公益研究计划项目 (LGG19F030011)。

作者简介: 王宪保 (1977-), 男, 山东菏泽人, 博士后, 副教授, 主要从事模式识别、神经网络及图像处理方向的研究。

姚明海 (1963-), 男, 浙江嘉兴人, 教授, 主要从事模式识别和图像识别方向的研究。

(YOLO, you only look once) 算法启发, 这类算法精度不及 Faster RCNN 这类基于候选区域方法, 但其检测速度却是前者的 10 倍, 后续又提出了 (SSD, single shot multi-box detector)^[15]、(RefineNet, refinement neural network)^[16]、(YOLOv2, you only look once v2)^[17] 以及 (YOLOv3, you only look once v3)^[18] 等。

本文将 (Mask RCNN, mask region based convolutional neural network)^[19] 作为基本框架, 并针对具体目标对象进行改进, 建立了一种用于滤袋的有效检测方法。本文的主要工作有: 首先, 针对固定卷积中感受野形状固定, 不能灵活地覆盖目标对象, 降低了特征提取效果这一问题, 采用可变形卷积网络 (DCN, deformable convolutional networks)^[20-21] 改进固定卷积, 通过引入额外的偏移来提升空间采样能力。其次, 提出一种改进的 (Soft-NMS, soft non-maximum suppression) 方法, 对 Mask RCNN 的 RPN 网络输出增加一次筛选, 提升候选区域的质量, 减少候选区域的数量。最后, 利用本文提出的目标检测器在滤袋数据集上进行实验, 结果证明可以有效检测滤袋开口, 相较于基准方法高了 2.4 个百分点。

本文的组织结构如下: 首先在第一节对基础框架等技术进行介绍; 第二节对本文提出的检测器进行详细介绍; 第三节通过实验, 验证了本文提出的检测器在滤袋开口检测中的有效性, 并通过与基准检测器的对比, 展现了本文检测器的性能优势; 第四节对全文工作进行总结与展望。

1 相关工作

1.1 Mask RCNN

Mask RCNN 模型建立在 Faster RCNN 的基础上, 将 Faster RCNN 中 (ROI Pooling, region of interest pooling) 替换为 (ROI Align, region of interest align), 消除了二次量化的误差, 并在头部网络中增加掩码分支用于实现实例分割, 提升了对对象的检测精度。

1.1.1 网络结构

Mask RCNN 的结构如图 1 所示。模型包含用于特征提取的主干网络、候选区域生成的 RPN、ROI Align 以及头部网络。主干网络的选择主要有 (VGG, visual geometry group)^[22]、(ResNet, deep residual network)^[23] 等, 其将一系列卷积进行堆叠, 逐层提取图像中的语义信息, 形成特征映射图。

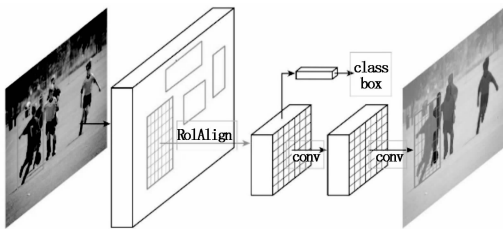


图 1 Mask RCNN 的网络结构^[10]

1.1.2 RPN

RPN 将任意大小的输入图像转换为一组含有分数的矩

形框作为 regionproposals, 并同后续网络共享特征提取图中的特征信息, 在输入在每个空间位置采样 k 个不同尺度的 anchors。通过后续二分类网络对这些 anchors 进行评估, 挑选出可能含有对象的 anchors, 并对其进行边界框的精修, 作为目标候选区域。

与目标对象交并比 (IOU, intersection over union) 最大的 anchors 以及和至少一个对象的 IOU 不小于预设阈值的 anchors 将被作为 proposal 层中的正样本, 其余作为负样本用以后续训练。

根据式 (1) 对 RPN 进行训练:

$$L_{rpn}(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (1)$$

式中, anchors 的索引值为 i , p_i 代表索引值为 i 的目标对象概率。 p_i^* 代表真实标签 (Ground Truth) 值 (1 表示正样本, 0 表示负样本)。 t_i 与 t_i^* 分别表示第 i 个 anchor 的边界修正数值与真实标签中的边界修正数值。 L_{cls} 是分类损失函数, 表示预测类别与标签之间的误差, 采用交叉熵损失函数实现。 L_{reg} 表示边界框回归的损失函数, 具体通过式 (2) 中的 SmoothL1 函数实现。 λ 是超参数, 用于分类损失与回归损失之间的平衡。总体损失通过梯度下降法修正。最终, 训练好的 RPN 可以生成高质量的候选区域为后续处理做准备。

$$SmoothL_1(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \quad (2)$$

1.1.3 ROI Align

ROI Align 是对 ROI Pooling 的优化。ROI Pooling 利用空间金字塔池化^[24], 将 RPN 的输出作为输入, 输出一组固定大小的特征张量作为后续处理的输入。但是 ROI Pooling 引入了两次量化操作, 带来了量化误差, 降低了 ROI 和其对应特征之间的一致性。针对此问题, ROI Align 通过双线性插值法解决, 对每个 ROI 先进行分割, 之后在分割得到的每个子区域中采样 K (K 一般取 4) 个点, 对每个点进行双线性插值得到其像素值, 之后将最大的像素值作为该子区域的像素值。这种方式避免了引入量化操作带来的量化误差, 实现了 ROI 与特征的一致, 提升了后续分类以及回归框的精度。

1.1.4 头部网络

由于引入了特征金字塔生成不同尺度的特征图, 所以需要根据 ROI 尺度选择特定层级的金字塔输出作为后续处理的输入, 具体选择如公式 (3):

$$k = k_0 + \log_2 \left(\frac{\sqrt{wh}}{224} \right) \quad (3)$$

式中, $k_0 = 5$, w, h 表示对应 ROI 区域的宽和高, 224 对应特征金字塔第五层的尺度。

经过 ROI Align 输出固定大小的特征张量, 经过两层全连接网络之后分别传送给分类器和回归器, 以实现目标对象的分类和边界框的回归。掩码分支的输入也由 ROI Align

产生, 但其尺寸大小与分类和回归不同。

网络训练采用多任务损失函数, 通过学习不断下降的损失函数的值, 直至获得最优解。损失函数的公式由式 (4) 给出。其中包含了 3 项, 分别对应分类损失, 边框回归损失以及掩码分割损失。

$$L = L_{cls} + L_{box} + L_{mask} \quad (4)$$

Mask RCNN 中使用固定卷积, 这类卷积在空间采样上不够灵活, 容易引入对象周边的噪声信息, 干扰后续的识别与分类。针对此问题, 本文引入可变形卷积来改进主干网络中的部分固定卷积, 以实现更加灵活的空间采样。针对 RPN 网络产生的候选区域提出一种改进的 Soft-NMS 方法进行进一步筛选与整合, 在降低候选框冗余的同时提升候选框的质量, 从而达到提升目标检测性能的目的。

2 融合可变形卷积的 Mask RCNN

本文将可变形卷积加入 Mask RCNN 中, 从而得到更好的空间采样结果, 并用改进的 Soft-NMS 进一步整合候选区域降低其冗余程度, 最终实现对滤袋开口的有效、精确检测。图 2 给出了本文目标检测模型的总体结构。整个模型分为 4 部分, 第一部分为主干网络, 实现图像特征的提取工作; 第二部分是 RPN, 用于候选目标的生成; 第三部分是 ROI Align, 它将候选区域对应的特征映射池化输出为固定大小的特征张量, 为后续处理做准备; 第四部分是头部网络, 由三个分支构成, 分别实现目标类别的识别、目标边界框的修正以及目标掩码信息的生成。最后将掩码信息与边界框信息整合得到更为精确的定位。

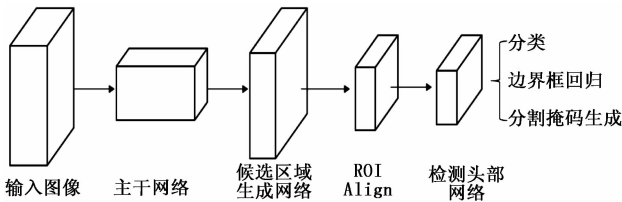


图 2 融合可变形卷积的 Mask RCNN 网络结构

2.1 主干网络

本文选取残差网络作为提取特征的主干网络。残差网络利用跳接 (skipconnect) 形成残差块, 通过堆叠残差块的方式构建残差网络。残差块可以用公式 (5) 表示:

$$y = F(x, \{W_i\}) + x \quad (5)$$

其中: y 表示残差块的输出, x 为残差块的输入, $F(x, \{W_i\})$ 表示待学习的残差映射。

残差网络通过跳接结构, 降低了随着模型深度的增加产生过拟合的可能性, 为构建更深的模型提供了基础。

传统的卷积结构空间采样不够灵活, 无法有效处理目标形变。本文引入可变形卷积对原始残差网络进行改进, 替换掉其高层中的固定卷积层, 从而更有效地实现对目标的特征提取, 提升检测效果。

传统的卷积从数学形式上可以用式 (6) 表示, 其中 p_n 代表卷积核中每个点对应中心点的空间偏移, R 表示卷积

核对应的各个空间位置偏移。

$$y(p_0) = \sum_{p_n \in R} w(p_n) x(p_0 + p_n) \quad (6)$$

相较于传统的卷积, 可变形卷积引入了一个空间偏移 Δp_n , 该偏移参数可以通过学习的方式获得。可变形卷积的表示如公式 (7):

$$y(p_0) = \sum_{p_n \in R} w(p_n) x(p_0 + p_n + \Delta p_n) \quad (7)$$

由于增加的 Δp_n 一般是小数, 并没有与之对应的输入, 所以通过双线性插值的办法来获得这一位置对应的输入。

随着卷积网络的加深, 顶层感受野的大小增加, 高层语义信息变得更多, 更适合分类, 但同时也损失底层信息, 这对小目标的检测会造成较大影响。因此本文构建特征金字塔 (FPN, feature pyramid network)^[25] 来融合高低层级的语义信息, 实现对各个层级卷积输出的捕获, 提升对小目标的检测效果。将残差网络中 $conv2, conv3, conv4, conv5$ 的输出分别用 C_2, C_3, C_4, C_5 表示。FPN 在 C_5 的输出上通过一个 1×1 的卷积, 将其通道数由原先的 2048 降为 256, 获得特征映射图 CP_5 。之后对 CP_5 进行上采样, 同时对 C_4 的输出经过一个 1×1 的卷积将其通道数下降为 256, 使用特征图相加的方式将两者结合得到 CP_4 。通过上述方式依次得到 CP_3, CP_2 , 之后对 CP_2, CP_3, CP_4, CP_5 分别使用一个 3×3 的卷积以降低上采样带来的混叠现象, 得到 P_2, P_3, P_4, P_5 , 构成了特征金字塔。

2.2 区域提议生成

候选目标区域的质量与数目直接影响到目标检测的效果与速度。本文通过 RPN 方法利用 anchor 机制获得原始的候选区域, 同时依据所用数据集中目标的尺度对 anchor 的大小做出了调整, 调整后的参数为 (32, 64, 128, 256), 长宽比为 (1, 0.5, 2)。

RPN 生成的原始提议框数目众多, 在这些候选框中有些包含背景图像, 为减少这类样本, 提升候选框的质量, 一般采用非极大值抑制对候选框进行筛选。由于 PRN 分类结果与候选框位置之间没有明确的线性关系, 所以传统 NMS 方法会导致很多虽然重叠但是包含目标不一样的候选框被直接删除。为此, 本文提出一种改进的 Soft-NMS 方法, Algorithm1 给出了改进后的伪代码。

Algorithm 1: 改进的 Soft-NMS 方法

Input: $B = \{b_1, \dots, b_N\}, S = \{s_1, \dots, s_N\}, G_t$, 其中 B 是一系列候选框, S 是其对应的得分, G_t 为 IOU 阈值

Output: $O = \{o_1, \dots, o_m\}, S = \{s_1, \dots, s_m\}, O$ 为输出框, S 为其对应的得分

Begin:

$O \leftarrow \emptyset$

While $B \neq \emptyset$ do:

$k \leftarrow \arg\max\{S\}$

$K \leftarrow b_k$

$L \leftarrow K$

$B \leftarrow B - K$

For b_i in B :

```

If IOU(K, bi) > Gt
L ← L ∪ bi
End if
si ← si · f(IOU(K, bi))
End for
K' ← f2(L)
O ← O ∪ K'
End while
    
```

上述算法中，计算得分的公式由式 (8) 给出，最佳候选框位置调整的方法由式 (9) 给出：

$$s_i = s_i e^{-\frac{iou(K, b_i)^2}{\sigma}} \quad (8)$$

$$K' = \sum_i^k \frac{Score(b_i)}{\sum_j Score(b_j)} b_i \quad (9)$$

式 (9) 中， k 代表了同当前最高得分的候选框之间 IOU 大于阈值的候选框的总数，之后根据得分进行加权修正。

2.3 网络的联合训练

本文采取端到端的训练方式，将 RPN 网络的训练任务同主干网络以及头部检测网络的检测任务结合起来，通过构建一个多任务损失函数，训练整体结构。损失函数由式 (10) 给出：

$$L_{total} = L_{cls} + L_{box} + L_{mask} + L_{rpn} \quad (10)$$

主干网络部分采用在 ImageNet 上训练的 ResNet50 + DCN 模型进行初始化，RPN 以及检测头采用在 COCO 数据集上训练的参数进行初始化。本文采用梯度下降法进行神经元参数的迭代调整。

3 实验结果与分析

本文使用本实验室构建的滤袋数据集。为增加训练样本个数，通过旋转、裁剪以及翻转等数据增强方式对数据集进行了增广。使用 labelImg 工具对数据集进行标注。得到数据集 1 和数据集 2，两个数据集共有样本 1 000 张。数据集 1 中包含 8 类不同圆形开口的滤袋，每类均有 100 张图像，按 8: 2 的比例分为训练集和测试集。数据集 2 中含有 2 类不同椭圆形开口的滤袋，每类均有 100 张图像，同样以 8: 2 的比例分为训练集和测试集。表 1 和表 2 分别给出了数据集 1 和数据集 2 的具体信息。图 3 和图 4 展示了数据集对应的部分图像。

表 1 数据集 1 中每一类包含的图片数目

类别	圆型小开口	圆型大开口	带十字花纹开口	带字花纹开口	带花纹开口	圆形中等开口	带十字开口	带横线开口
数目	100	100	100	100	100	100	100	100

表 2 数据集 2 中每一类包含的图片数目

类别	紫色椭圆	白色椭圆
数目	100	100

3.1 评价标准

本文的评价指标为各类平均精度 (mAP, mean average

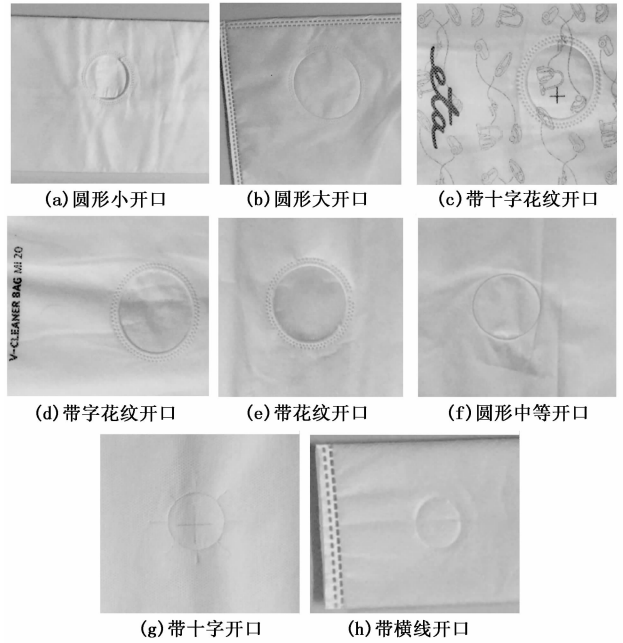


图 3 数据集 1 图片示例

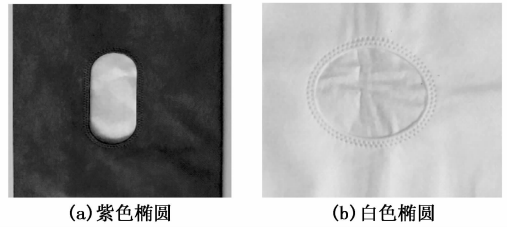


图 4 数据集 2 图片示例

precision)，中心偏移距离以及检测速度。本文对比的模型为 Faster RCNN 以及 Mask RCNN 两种主流的两阶段目标检测模型。

box_{gt} 为真实值标签用 $(y1_{gt}, x1_{gt}, y2_{gt}, x2_{gt})$ 表示，预测框 box_{pred} 用 $(y1_{pred}, x1_{pred}, y2_{pred}, x2_{pred})$ 表示。两者之间的距离记作 $d = dis(box_{gt}, box_{pred})$ ，可以根据公式 (11) 计算：

$$d = dis(box_{gt}, box_{pred}) = \sqrt{(y_{gt} - y_{pred})^2 + (x_{gt} - x_{pred})^2} \quad (11)$$

其中： $y_{gt} = (y1_{gt} + y2_{gt})/2, x_{gt} = (x1_{gt} + x2_{gt})/2, y_{pred} = (y1_{pred} + y2_{pred})/2, x_{pred} = (x1_{pred} + x2_{pred})/2$ 。

本文的模型在数据集 1 与数据集 2 上进行评估。

3.2 参数设置

实验的硬件环境为 Intel Corei5-9400 CPU, NVIDIA GTX 1080 GPU。软件环境为 Tensorflow 架构，实验语言为 Python。训练时设置学习率为 0.001，权重衰减系数为 0.000 1，采用带动量的 (SGD, stochastic gradient descent) 优化器迭代训练，动量设置为 0.9，训练迭代 20 000 次，batch-size 设置为 2。实验主要由以下几部分组成：第一个实验将对比本文算法与基准算法之间的性能差异；第二个实验为消融实验将验证提出的改进 Soft-NMS 方法的

有效性; 第三个实验将数据集 1 和数据集 2 合并, 检验算法在输入类别增加情况下的鲁棒性。

3.3 算法比较

实验一: 在本实验中, 通过将 Faster RCNN、Mask RCNN 以及本文的融合可变形卷积 Mask RCNN 在滤袋数据集 1 上训练和测试, 得到如表 3 展示的结果。图 6 给出了部分检测结果。

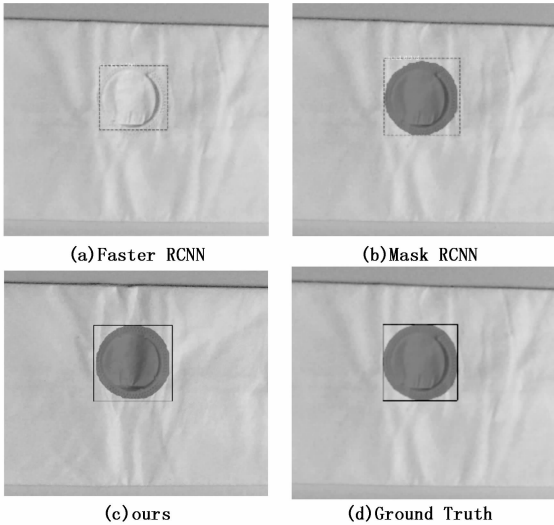


图 6 在数据集 1 上检测的结果 (示例)

表 3 在吸尘器尘袋测试集上的检测准确率以及中心偏移距离

	Backbone	mAP	中心偏移距离	Fps
Faster RCNN	VGG16	88.1	24.34	1.12
Faster RCNN	ResNet50	90.6	20.62	1.23
Mask RCNN	ResNet50	93.2	13.72	0.95
Ours	ResNet50+DCN	95.6	11.76	0.93

在表 3 中可以看到, 本文的算法在中心偏移距离这一指标上优于对比算法, 相较于 Faster RCNN 提升了 11.58 个像素。Soft-NMS 过程对候选框的重整提高了这一指标。

本文构建的特征金字塔实现了特征的融合, 提升了对小目标的检测精度, 使得模型的漏报率下降, 提升了模型的召回率。藉此, 总体的 mAP 相较于没有特征金字塔结构的 Faster RCNN 也得到了提升。同时, 可变形卷积带来了更加灵活的空间采样, 引入改进 Soft-NMS 的区域提议生成网络提供了更高质量的候选框, 给本文方法在 mAP 上带来了高于对比算法的性能, 相较于 Mask RCNN 提升了 2.4 个百分点。

在检测速度方面, 由于本文模型基于 Mask RCNN 构建, 其增加了掩码分支且本文在模型中加入金字塔结构以及可变形卷积, 在一定程度上提升了模型的复杂度, 致使本文算法相较于 Faster RCNN 的推理速度有所下降, 但是和 Mask RCNN 相当。

实验 2: 为了验证改进 Soft-NMS 的有效性在实验数据集 1 上进行消融实验。实验结果如表 4 所示。对比采用的是 NMS 方法, 本文使用的是改进的 Soft-NMS 方法, 其

余部分两者相同皆为融合可变形卷积的 Mask RCNN。可以看到本文提出的 Soft-NMS 方法相较于传统 NMS 在数据集 1 上 mAP 指标实现了 1.2 个百分点的提升, 同时中心偏移距离提升了 0.36 个像素, 验证了本文提出的改进 Soft-NMS 方法的有效性, 经过重整的候选框质量确实优于原始候选框。

实验 3: 为了验证算法的泛化性能, 本文用数据集 1 和数据集 2 对算法进行验证, 结果如表 5 所示。可以看到引入可变形卷积并使用改进 Soft-NMS 方法后, 本文模型性能依旧高于其他对比算法, 证明了本文方法在增加对象类别之后仍然可以具有较好性能。为了进一步检验其性能, 将来自不同数据集的对象放在同一环境采集, 继续测试, 结果如图 7 所示。实验结果证明本文模型可以有效地检测不同滤袋开口, 具有良好的鲁棒性。

表 4 改进 Soft-NMS 方法的有效性

Method	mAP	中心偏移距离	Fps
NMS	94.4	12.12	0.94
Improved Soft-NMS	95.6	11.76	0.93

表 5 在数据集 1+2 上的检测结果

	Backbone	mAP	中心偏移距离	Fps
Faster RCNN	ResNet50	89.5	22.63	1.17
Mask RCNN	ResNet50	92.0	15.26	0.94
ours	ResNet50+DCN	94.5	12.46	0.92

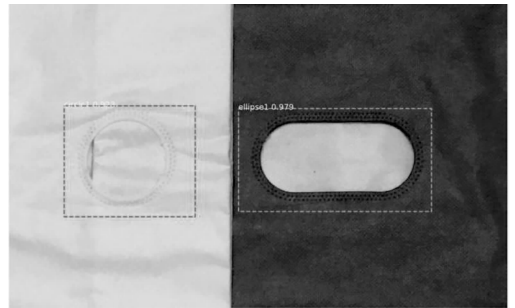


图 7 在同一图片中两种不同尘袋的检测效果

4 结束语

本文提出了一种基于深度学习的滤袋开口检测算法, 实验结果表明, 本文算法在滤袋开口检测问题上比传统基于深度学习的目标检测方法表现得更好, 且检测速度较快。由于本文构建的数据集规模较小, 后续工作中将尝试建立种类更为全面的大型数据集供滤袋开口检测算法的训练和测试, 并且进一步通过剪枝算法压缩模型体积, 提升运算速度。

参考文献:

- [1] 田合雷, 丁 胜, 于长伟, 等. 基于目标检测及跟踪的视频摘要技术研究 [J]. 计算机科学, 2016, 43 (11): 297-299.
- [2] Schmidhuber J. Deep learning in neural networks: an overview [J]. Neural Network, 2015, 61 (8): 85-117.

[3] 许必宵, 宫 婧, 孙知信. 基于卷积神经网络的目标检测模型综述 [J]. 计算机技术与发展, 2019, 11: 87-92.

[4] Dalal N, Triggs B. Histograms of oriented gradients for human detection [A]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) [C]. IEEE, 2005.

[5] Lowe D. Distinctive image features from scale-invariant key points [J]. Computer Vision, 2003, 20: 91-110.

[6] Felzenszwalb P, Mcallester D, Ramanan D. A discriminatively trained, multiscale, deformable part model [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. 2008, 8: 1-8.

[7] Freund Y, Schapire R E. Experiments with a new boosting algorithm [A]. Proceedings of Thirteenth International Conference on Machine Learning [C]. 1996: 148-156.

[8] Dalal N, Triggs B. Histograms of oriented gradients for human detection [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]. 2005: 886-893.

[9] Girshick R. Fast R-CNN [A]. IEEE International conference on computer vision [C]. New York: IEEE Computer Society, 2015: 1440-1448.

[10] Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [A]. International Conference on Neural Information Processing Systems [C]. Massachusetts: MIT Press, 2015: 91-99.

[11] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39 (4): 640-651.

[12] Uijlings J R R, van de Sande K E A. Selective search for object recognition [J]. Computer Vision, 2013, 104 (2): 154-171.

[13] Zitnick C L, Dollar P. Edge boxes: locating object proposals from edges [M]. ECCV. Springer International Publishing, 2014.

[14] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [A]. Computer Vision and Pattern Recognition [C]. New York: IEEE Press, 2016: 779-788.

[15] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multi-

box detector [A]. European Conference on Computer Vision [C]. 2016: 21-37.

[16] Lin G, Milan A, Shen C, et al. RefineNet: multi-path refinement networks for high-resolution semantic segmentation [A]. The IEEE Conference on Computer Vision and Pattern Recognition [C]. 2017: 1925-1934.

[17] Redmon J, Farhadi A. YOLO9000: better, faster, stronger [A]. The IEEE Conference on Computer Vision and Pattern Recognition [C]. 2017: 7263-7271.

[18] Redmon J, Farhadi A. YOLOv3: An incremental improvement [J/OL]. arXiv, [2018-04-8]. <https://arxiv.org/abs/1904.02767>.

[19] He K, Gkioxari G, Dollar P, et al. Mask R-CNN [A]. 2017 IEEE International Conference on Computer Vision (ICCV) [C]. IEEE Computer Society, 2017.

[20] Dai J F, Qi H Z, Xiong Y W, et al. Deformable convolutional networks [A]. IEEE International Conference on Computer Vision [C]. 2017: 764-773.

[21] Zhu X Z, Hu H, Lin S, et al. Deformable ConvNets v2: more deformable, better results [A]. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2019: 9308-9316.

[22] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [A]. International Conference on Learning Representations [C]. 2015.

[23] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [A]. IEEE Conference on Computer Vision & Pattern Recognition [C]. IEEE Computer Society, 2016: 770-778.

[24] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37 (9): 1904-1916.

[25] Lin T Y, Dollar P, Girshick R, et al. Feature pyramid networks for object detection [C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2017: 2117-2125.

（上接第 20 页）

[5] 张文瑶. 对新时期海军舰船装备维修策略的思考 [J]. 中国修船, 2010, 23 (4): 45-49.

[6] 马 剑, 吕 琛, 陶来发. 船舶主推进系统故障预测与健康管理设计 [J]. 南京航空航天大学学报, 2011 (s1): 119-124.

[7] 姚晓山, 张卫东, 周 平, 等. 基于油液监测的船舶柴油机故障预测与健康管理系统研究 [J]. 武汉理工大学学报: 交通科学与工程院, 2014, (38) 4: 874-877.

[8] 谢 娜, 雷江妮. 民机故障预测与健康管理系统顶层架构设计技术研究 [J]. 计算机测量与控制, 2020, 28 (2): 19-22.

[9] 姚玉玲, 高小红, 王 恒. 航空电子设备故障预测与健康管理系统 (PHM) 技术浅谈 [A]. 探索 创新 交流——第六届中国航空学会青年科技论坛 [C]. 2014: 1471-1476.

[10] 赵媛媛, 陈 妹. 国外军用车辆故障预测与健康管理系统的发展 [J]. 国外坦克, 2017 (11): 38-41.

[11] 邓晓衡, 关培源, 王志文. 基于综合信任的边缘计算资源协同研究 [J]. 计算机研究与发展, 2018, 55 (3): 5-33.

[12] 边缘计算产业联盟. 边缘计算参考架构 2.0 (下) [P]. 自动化博览, 2018 (3): 58-61.

[13] 施巍松, 张星洲, 王一帆. 边缘计算: 现状与展望 [J]. 计算机研究与发展, 2019, 56 (1): 73-93.

[14] 赵梓铭, 刘 芳, 蔡志平, 等. 边缘计算: 平台、应用与挑战 [J]. 计算机研究与发展, 2018, 55 (2): 327-337.

[15] 陈 洁. 边缘计算下无线通信接入认证技术研究 [D]. 成都: 电子科技大学, 2019.