

基于YOLOv7的垃圾焚烧分类检测研究

谢能勇

安徽理工大学经济与管理学院, 安徽 淮南

收稿日期: 2023年1月22日; 录用日期: 2023年2月17日; 发布日期: 2023年2月27日

摘要

为实现自动垃圾检测和抓取工作, 解决垃圾站操作工人的不准确性和劳动力, 本文使用基于YOLOv7算法的实时垃圾检测算法, 可以通过垃圾池区域中的摄像头对各分区的垃圾进行实时拍照检测, 通过吊车夹爪对待抓取的垃圾进行实时抓取进入待焚烧区域, 有效提高生产效率。实验结果表明, 本文算法在训练100轮次的mAP值为87.7%, 可较好地满足实际垃圾实时工业处理的需求。

关键词

目标检测, YOLOv7, 实时监控

Research on Classification and Detection of Waste Incineration Based on YOLOv7

Nengyong Xie

School of Economics and Management, Anhui University of Science and Technology, Huainan Anhui

Received: Jan. 22nd, 2023; accepted: Feb. 17th, 2023; published: Feb. 27th, 2023

Abstract

In order to realize automatic garbage detection and grasping work and solve the inaccuracy and labor of the workers operating the refuse collection point, this paper uses a real-time garbage detection algorithm based on the YOLOv7 algorithm, which can detect the garbage in each partition by taking real-time photos through the camera in the garbage pool area and grasp the garbage to be grasped into the area to be incinerated in real-time through the crane jaws, effectively improving the production efficiency. The experimental results show that the algorithm of this paper has a mAP value of 87.7% in 100 rounds of training, which can better meet the needs of real garbage in real-time industrial processing.

Keywords

Target Detection, YOLOv7, Real-Time Monitoring

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

目前, 传统垃圾焚烧和回收主要操作是通过垃圾回收站的操作工人对垃圾进行手动分拣及垃圾抓取至待焚烧区域。这样不仅增加了劳动成本, 同时操作工人需要凭借经验判断, 不利于垃圾充分燃烧, 因此, 本文采用一种基于深度学习目标检测的 YOLOv7 算法解决了传统垃圾检测存在的问题。

目标检测(Object Detection)作为计算机视觉领域的核心问题之一, 其任务是在图像或图像序列中对目标进行精确定位和分类。但目标检测任务中存在各类干扰因素, 如目标物体本身的纹理、形态等, 以及成像时的光照强弱、遮挡和模糊等。基于上述因素, 目标检测性能成为计算机视觉领域亟待解决的问题, 针对这些因素, 众多学者对目标检测算法进行了研究与改进。One-Stage 目标检测算法不单独选取候选区域, 省略了候选区域生成步骤, 将特征提取、目标分类和位置回归整合到一个阶段进行操作, 代表算法有 YOLO 系列、SSD 系列、RetinaNet、CornerNet、EfficientDet 等[1], 下面将分别进行介绍和分析。Two-Stage 目标检测算法先对图像进行选取候选区域操作, 再对候选区域进行分类和位置定位, 使检测过程划分为两个阶段, 代表算法有 R-CNN (Regions Convolutional Neural Network)、Fast R-CNN、Faster R-CNN、Cascade R-CNN 等, 下面将分别进行介绍和分析。

针对以上问题, 本文提出了一种基于 YOLOv7 的目标检测算法, 针对实时垃圾检测及抓取算法。具体工作包括如下: 1) 自建一个关于实时垃圾及吊车夹爪数据集; 2) 基于 YOLOv7 算法对垃圾和夹爪进行实时检测。

2. YOLOv7 的网络结构

2.1. 扩展的高效层聚合网络

在大多数关于设计高效架构的文献中, 主要考虑的因素不外乎是参数的数量、计算量和计算密度。Bello 等人[2]从内存访问成本的特点出发, 分析了输入/输出通道比、架构的分支数和元素的操作对网络的影响、架构的分支数量以及对网络推理速度的逐元操作。Bochkovskiy 等人[3]在进行模型扩展时, 还考虑了激活问题, 即更多地考虑卷积层输出张量中的元素数量。CSPVoVNet 的设计图 1 中的 CSPVoVNet [4]的设计是 VoVNet [5]的一个变体。VoVNet [4]的变化, 除了考虑前述的基本设计问题、基本的设计关注点外, CSPVoVNet 的架构[6]还分析了梯度路径, 以使不同层的权重来学习更多不同的特征。上述的梯度分析方法使推断更快、更准确。图 1 中的 ELAN [7], 考虑了以下设计策略——“如何设计一个高效的网络? ”。他们得出了一个结论: 通过控制最短的最长梯度路径, 一个更深的网络可以有效地学习和收敛。在本文中, 我们提出了基于 ELAN 的扩展 ELAN (E-ELAN), 其主要结构如图 1 所示。

无论梯度路径长度和堆积大规模 ELAN 中计算块的数量, 它都达到了一个稳定的状态。如果更多的计算块被无限制地堆叠, 这种稳定状态可能会被破坏, 参数的利用率也会下降。所提出的 E-ELAN 通过

使用扩展、移动、合并基数实现了在不破坏原始梯度路径的情况下不断提高网络的学习能力。我们的策略是利用群卷积来扩展计算块的通道和基数。在结构方面，E-ELAN 只改变了计算块的结构，而过渡层的结构完全没有改变。我们的策略是使用组卷积来扩展计算块的通道和基数，同时应用相同的组参数和信道倍增器用于计算层的所有计算块。此时，每一组特征图中的通道数将与原始体系结构中的通道数相同。然后，将每个计算块计算出的特征图根据设置的组参数 g 整分为 g 组，然后将它们连接在一起。最后，我们添加了 g 组特征映射来执行合并基数。除了维护原始的 ELAN 设计架构外，E-ELAN 还可以指导不同的计算块组来学习更多样化的特性。

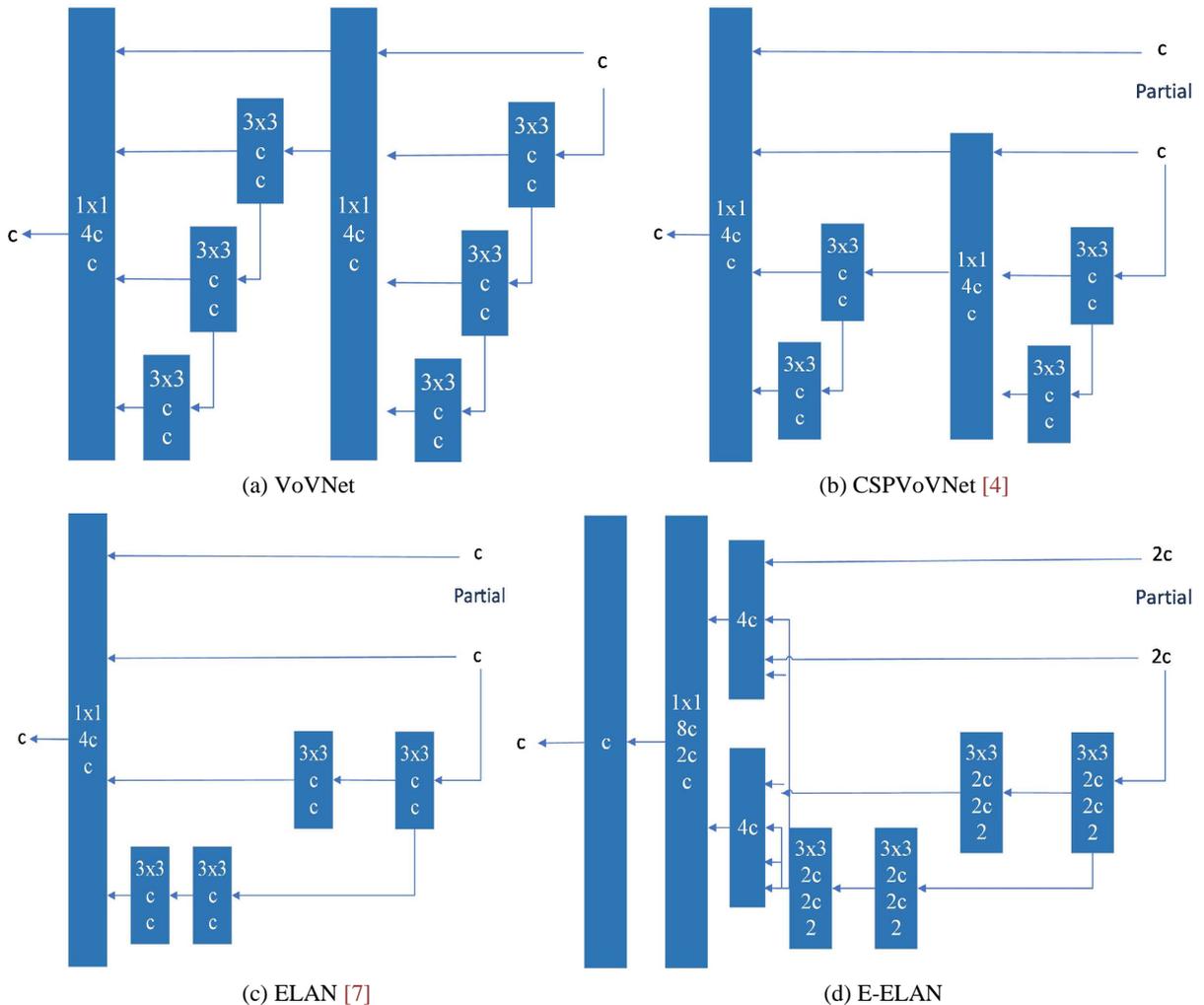


Figure 1. Extended efficient layer aggregation network. The proposed Extended ELAN (E-ELAN) does not change the gradient transfer path of the original architecture at all, but uses group convolution to increase the cardinality of the added features, and combines the characteristics of different groups in a way of shuffling and merging cards. This way of operation can be enhanced by different feature maps and improve the use of parameters and computations

图 1. 扩展的高效层聚合网络。所提出的扩展 ELAN (E-ELAN) 完全不改变原始架构的梯度传输路径，而是使用组卷积来增加所增加的特征的 cardinality，并将不同组的特征以洗牌和合并卡度的方式进行组合。这种操作方式可以增强由不同的特征图，并改善参数和计算的使用

2.2. 基于串联的模型缩放

模型缩放的主要目的是调整模型的一些属性，生成不同尺度的模型以满足不同推理速度的需要。例

如, EfficientNet 网络[8]的缩放模型考虑了宽度、深度和分辨率。至于规模化的 YOLOv4, 其缩放模型是调整阶段的数量。在[3]中, Bochkovskiy 等人分析了原始卷积和组卷积对进行宽度和深度缩放时的参数和计算量的影响, 并设计了相应的模型尺度方法, 用它来设计相应的模型缩放方法。

上述方法主要用于诸如 PlainNet 或 ResNet 等架构中。当这些架构在执行放大或缩小过程时, 每一层的内度和出度都不会发生变化, 因此, 我们可以独立分析每个缩放因子对参数量和计算量的影响。然而, 如果这些方法应用于基于连接的架构, 我们会发现当扩大或缩小执行深度, 基于连接的翻译层的计算块将减少或增加, 如图 2 所示。

从上述现象可以推断, 我们不能对基于连接的模型分别分析不同的比例因子, 而必须一起考虑。以放大深度为例, 这样的动作会导致过渡层的输入通道和输出通道之间的比率变化, 从而导致模型的硬件使用量的减少。因此, 我们必须为一个基于连接的模型提出相应的复合模型缩放方法。当我们缩放一个计算块的深度因子时, 我们还必须计算该块的输出通道的变化。然后, 我们将对过渡层以相同的变化量进行宽度因子缩放, 结果如图 2。所示我们提出的复合尺度方法可以保持模型在初始设计时的特性, 并保持最优结构。

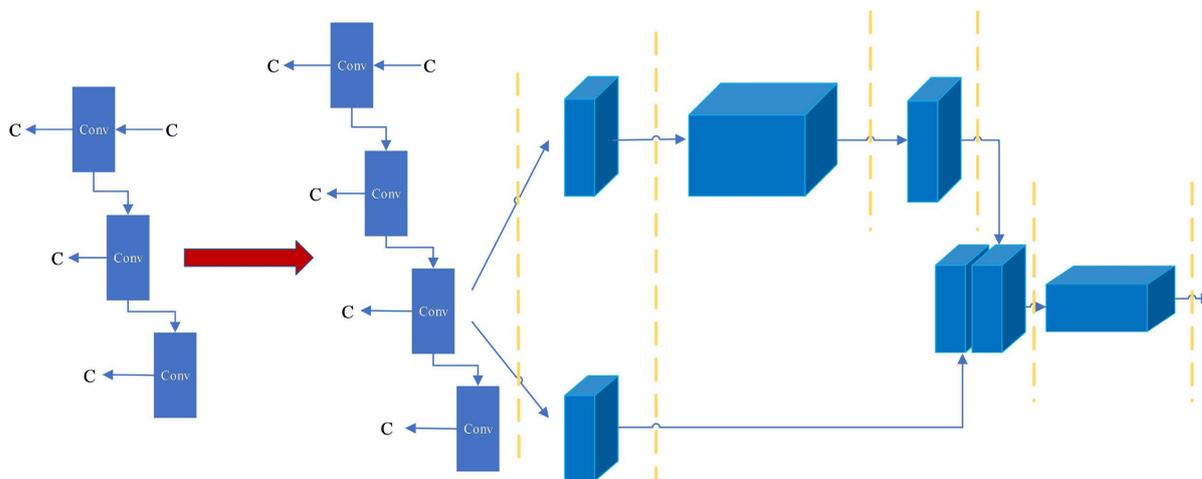


Figure 2. Model scaling for connection-based models. We observe that the output width of the computational block also increases when depth scaling is performed on the connection-based model. This phenomenon causes the input width of subsequent transport layers to increase. Therefore, we propose that when model scaling is performed on a connection-based model, it is only necessary to adjust the depth in the computation block and perform corresponding width scaling on the remaining transport layers

图 2. 基于连接的模型的模型缩放。我们观察到, 当在基于连接的模型上进行深度缩放时, 计算块的输出宽度也会增加。这种现象会导致后续传输层的输入宽度增加。因此, 我们提出了在基于连接的模型上进行模型缩放时, 只需要调整计算块中的深度, 并对剩余的传输层进行相应的宽度缩放

3. 训练技巧

3.1. 重新设计模型参数化

尽管 RepConv [9]在 VGG [10]上取得了优异的性能, 但当我们已将它直接应用于 ResNet [11]、DenseNet [8]和其他架构时, 它的精度将显著降低。我们使用梯度流传播路径来分析重新参数化的卷积应该如何与不同的网络相结合, 还相应地设计了计划中的重新参数化的卷积。RepConv 实际上结合了 3×3 卷积、 1×1 卷积和在一个卷积层中的身份连接。通过分析 RepConv 与不同架构的组合及其性能, 我们发现, RepConv 中的身份连接破坏了 ResNet 中的残差和 DenseNet 中的连接, 为不同的特征图提供了更多的梯度多样性。基于上述原因, 我们使用没有身份连接的 RepConv (RepConvN)来设计计划

中的重新参数化卷积的体系结构。在我们的思维中，当具有残余或连接的卷积层被重新参数化的卷积所取代时，不应该存在身份连接。图 3 显示了我们在 PlainNet 和 ResNet 中使用的“计划重新参数化卷积”的一个示例。对于基于残差的模型和基于连接的模型中完整计划的再参数化卷积实验，它将在消融研究环节中提出。

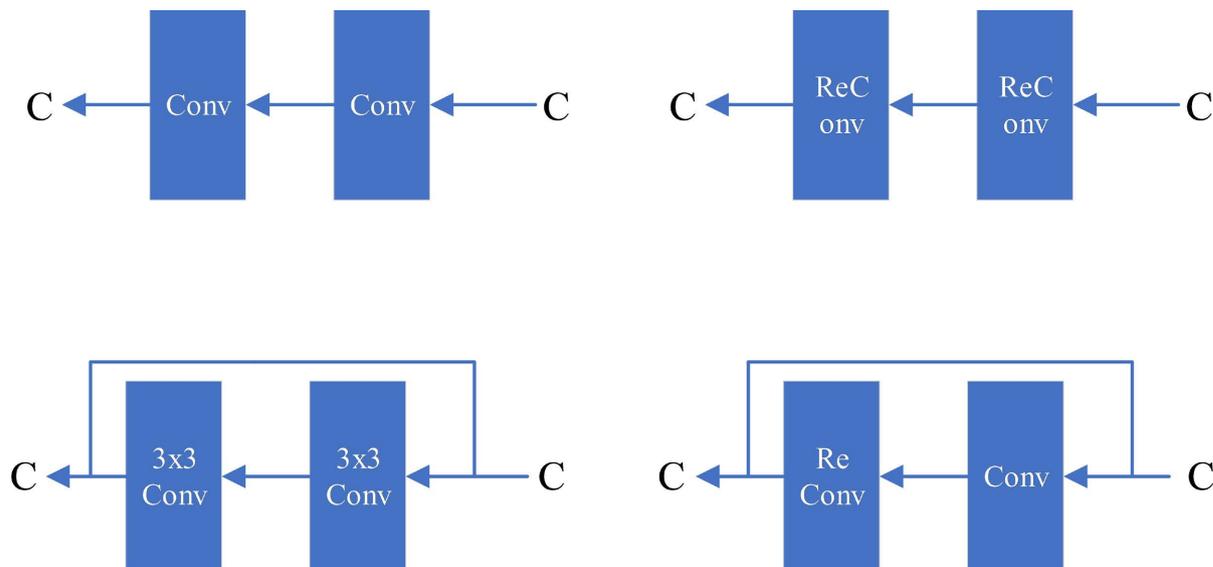


Figure 3. The planned reparameterized model. In the proposed planned reparameterization model, we find that a layer with residual or connected connections, its RepConv should not have identity connections. In these cases, it can be replaced by a RepConvN that does not contain an identity connection

图 3. 计划中的重新参数化模型。在提出的计划再参数化模型中，我们发现一个具有残余或连接连接的层，它的 RepConv 不应该具有身份连接。在这些情况下，它可以被不包含身份连接的 RepConvN 替换

3.2. 粗标签用于辅助头、细标签用于引导头

深度监督[10]是一种常用于训练深度网络的技术。其主要概念是在网络的中间层增加额外的辅助磁头，以及以辅助损失为导向的浅层网络权值。即使对于像 ResNet [11]和 DenseNet [1]这样通常收敛得很好的体系结构，深度监督仍然可以显著提高模型在许多任务上的性能。图 4 分别显示了“没有”和“有”深度监督的目标检测器架构。在本文中，我们将负责最终输出的头称为引导头，将用于辅助训练的头称为辅助头。

接下来，我们想讨论一下标签分配的问题。过去，在深度网络的训练中，标签分配通常直接指地面真相，并根据给定的规则生成硬标签。然而，近年来，如果我们以目标检测为例，研究者经常利用网络预测输出的质量和分布，然后结合地面真相考虑，使用一些计算和优化方法来生成可靠的软标签。例如，YOLO [12]使用边界盒回归预测和地面真实的 IoU 作为客观性的软标签。在本文中，我们将网络预测结果与地面真相一起考虑，然后将软标签分配为“标签分配者”的机制。

无论辅助头或领导头的情况如何，都需要对目标进行深度监督培训。在软标签分配人相关技术的开发过程中，我们偶然发现了一个新的衍生问题，即“如何将软标签分配给辅助头和铅头？”，据我们所知，相关文献迄今尚未对这一问题进行探讨。目前最常用的方法的结果如图 4 所示，即将辅助头和铅头分离，然后利用它们自己的预测结果去执行标签分配。本文提出的方法是一种新的标签分配方法，通过铅头预测来引导辅助头和铅头。换句话说，我们使用铅头预测作为指导，生成从粗到细的层次标签，分别用于辅助头和铅头的学习。所提出的两种深度监督标签分配策略分别如图 4 所示。

铅头引导标签分配器主要根据铅头的预测结果和地面真实值进行计算，并通过优化过程生成软标签。

这组软标签将作为辅助头和铅头的目标训练模型。这样做的原因是铅头具有相对较强的学习能力，因此，由此产生的软标签应该更能代表源数据与目标之间的分布和相关性。此外，我们还可以将这种学习看作是一种广义剩余学习。通过让较浅的辅助头直接学习铅头已经学习到的信息，铅头将更能专注于学习尚未学习到的残余信息。

从粗到细的铅头引导标签分配人也使用铅头的预测结果和地面真相来生成软标签。然而，在此过程中，我们生成了两组不同的软标签，即粗标签和细标签，其中，细标签与铅头引导标签分配者生成的软标签相同，粗标签是通过允许更多的网格被视为积极样本分配过程的约束而生成的。原因是一个辅助头的学习能力不是那么强大的领导头，为了避免丢失需要学习的信息，我们将专注于优化召回辅助头的目标检测任务。对于铅头的输出，我们可以从高查全率结果中过滤高精度结果作为最终输出。然而，我们必须注意到，如果粗标签的额外重量接近于对于精细标签，它可能在最终预测时产生不良的先验。因此，为了使这些超粗的正网格的影响更小，我们在解码器中进行了限制，从而使超粗的正网格不能完美地产生软标签。上述机制允许在学习过程中动态调整细标签和粗标签的重要性，并使细标签的可优化上界始终高于粗标签。

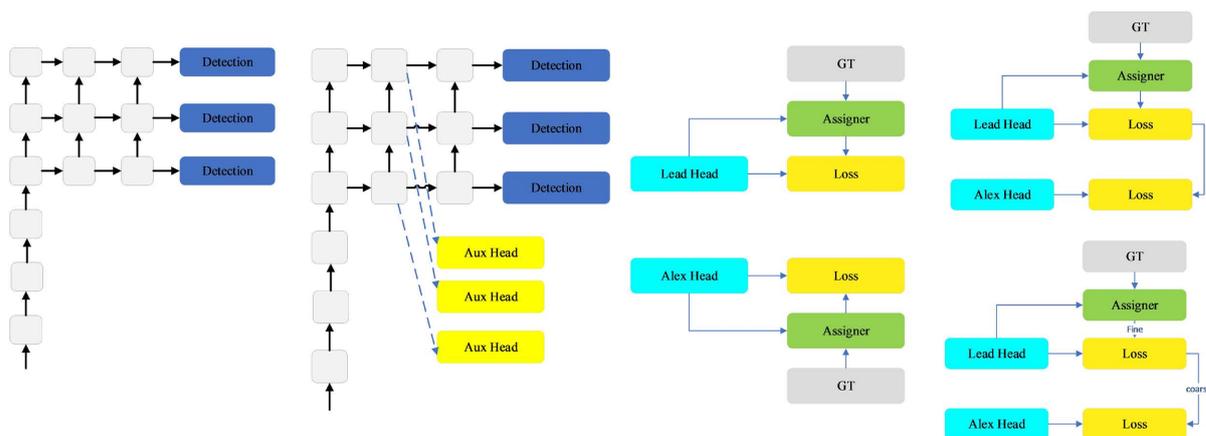


Figure 4. Auxiliary equipment is rough, lead-head label dispenser equipment is fine. Compared to the normal model, the schema in this figure has auxiliary headers. Different from the usual independent label assigners, we propose lead-guided label assigners and lead-guided label assigners from coarse to fine. With lead prediction and ground truth, the proposed label assigner is optimized, and the labels for both the training lead and the auxiliary head are obtained. The detailed coarse-to-fine implementation method and constraint design details will be elaborated in the attached appendix

图 4. 辅助设备设备粗糙，铅头标签分配器设备精细。与普通模型相比，此图中的图式有辅助头。与通常的独立标签分配人不同，我们提出了铅头引导标签分配人和从粗到细的铅头引导标签分配人。通过铅头预测和地面真相，对所提出的标签分配人进行优化，同时得到训练铅头和辅助头的标签。详细的从粗到细的实现方法和约束设计细节将在附件附录中详细阐述

3.3. 其他免费训练包

在本节中，我们将列出一些可训练的包。这些免费赠品是我们在训练中使用的一些技巧，但最初的概念并不是由我们提出的。这些免费赠品的训练细节将在附录中详细阐述，包括 1) conv-bn 激活拓扑中的批处理归一化：这部分主要将批处理归一化层直接连接到卷积层。其目的是将推理阶段批量归一化的均值和方差整合到卷积层的偏差和权重中。2) YOLOR [12]中的隐式知识结合卷积特征映射和乘法方式：YOLOR 中的隐式知识可以通过推理阶段的预计算简化为向量。这个向量可以与之前或后续卷积层的偏差和权重相结合。3) EMA 模型：EMA 是在平均策略[5]中使用的一种技术，在我们的系统中，我们纯粹使用 EMA 模型作为最终的推理模型。

4. YOLOv7 的垃圾检测算法

4.1. 数据集的构建

1) 数据获取

根据中国节能环保集团有限责任公司垃圾池区域操作工的反馈，垃圾池区域主要的目标物包括垃圾及吊车夹爪 2 种，以往需要工人手动操作抓取垃圾将其放入待烧区域，该工作不仅耗费人力、财力，同时，操作工人无法准确判断待烧区域的垃圾量。为了可以实现垃圾自动检测及实时抓取操作，本章节选取 2 种目标物进行检测，通过对垃圾池的监控摄像视频数据按照不同时间、不同区域、不同料口进行分帧截取，具体情况如图 5 所示。共计图片数据 2662 张，图像分辨率为 1920×1280 。

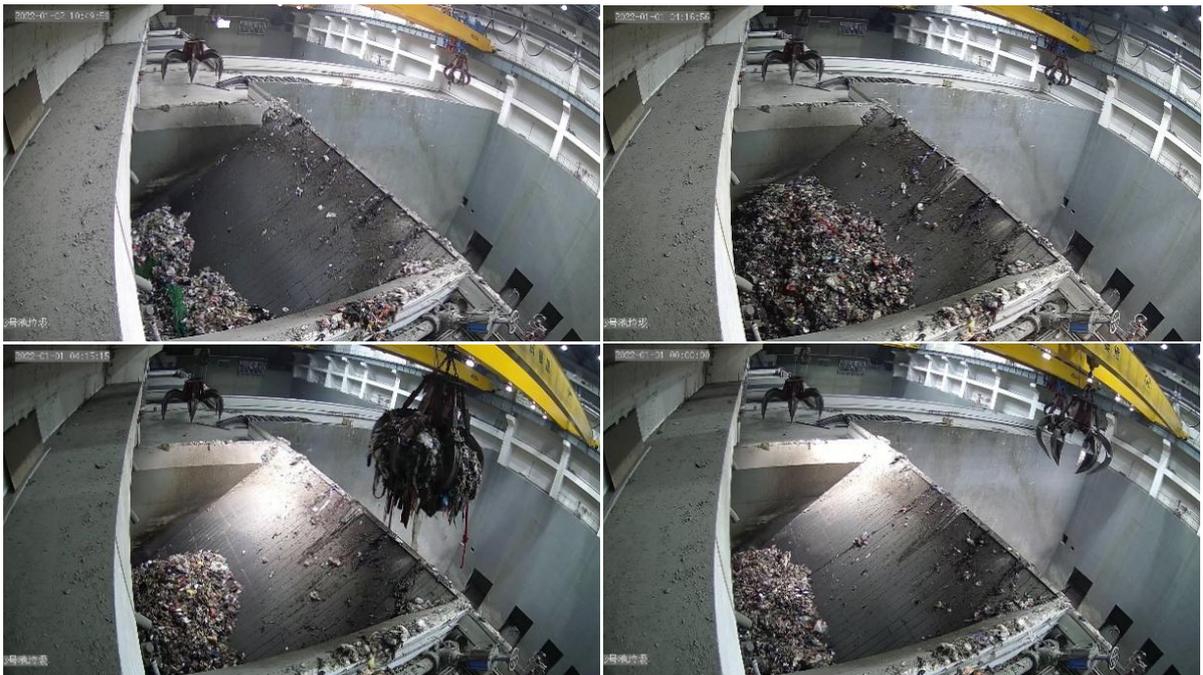


Figure 5. Schematic diagram of the garbage to be burned and the crane gripper

图 5. 待烧垃圾及吊车夹爪示意图

2) 数据标注

模型训练时输出预测值和真实值的误差通过损失函数表征，经过反向传播不断更新权重参数，寻找最优解，因此，在实验之前，需要进行数据标注工作，来获得真实标签。本文使用 LabelImg 工具对所有图片按 VOC2007 数据集标准格式进行标注，得到 xml 格式的标签文件，标注界面如图 6 所示。总共标注 2 种目标物：垃圾(waste)、夹爪(gripper)。

3) 数据划分

将数据集按照 9:1 的比例划分为训练集和验证集，其中，训练集 2136 张，验证集 526 张，具体数据集划分情况如表 1 所示。模型在训练时所有训练集和验证集全部输入给网络，模型测试时在测试集上对垃圾及夹爪进行分别评估。

4.2. 实验环境及评价准则

1) 实验环境

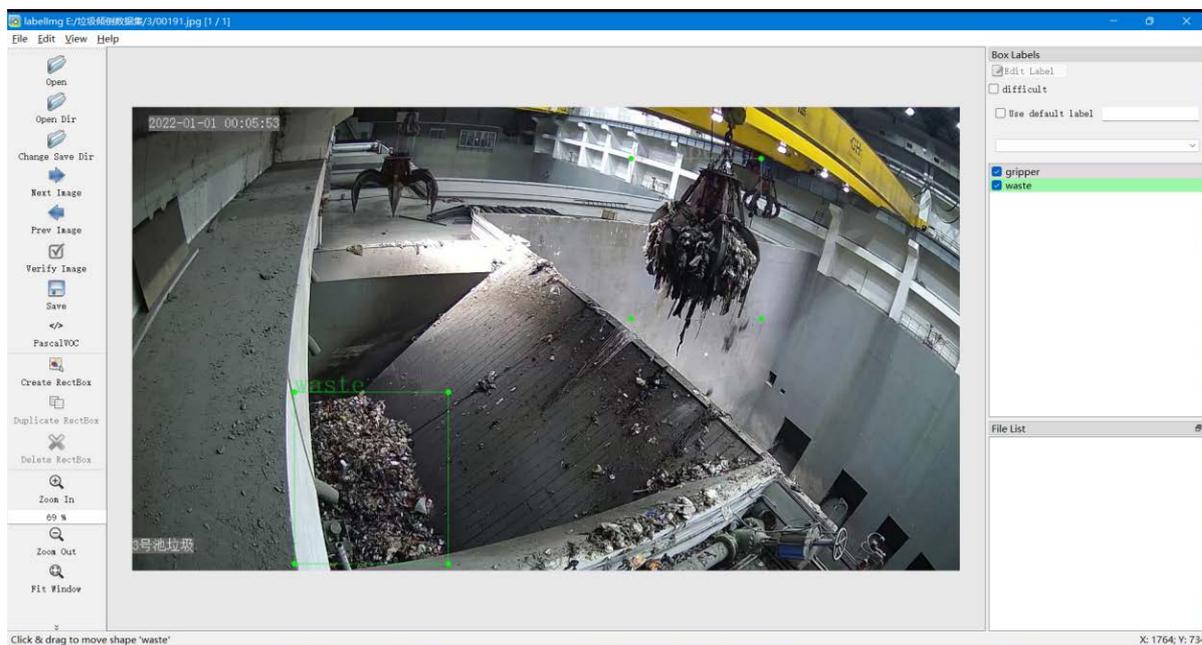


Figure 6. Schematic diagram of data annotation

图 6. 数据标注示意图

Table 1. Data set division details

表 1. 数据集划分详情

数据集类别	图片数量	垃圾堆目标个数	夹爪目标个数
训练集	2136	2727	4326
验证集	526	785	1124

本文全部实验的软硬件环境及相关版本参数为 Ubuntu 16.04.4 LTS 系统下，采用 Pytorch 深度学习框架搭建。计算机硬件配置包括 NVIDIA GeForce RTX 3090 24GB GPU、AMD Ryzen 9 5950X 16-Core CPU 和 16 GB 内存。

2) 评价准则

本文主要采用平均精度 AP (Average Precision)、平均精度均值 mAP (Mean Average Precision) 和 FPS (Frame Per Second) 作为评价指标来衡量算法性能。FPS 是衡量检测速度的最直观指标，指的是算法每秒可检测的图像帧数，在实际应用中，实时检测的标准通常为 $FPS \geq 30$ 。平均精度 AP 的计算涉及到精确率 (Precision) 和召回率 (Recall) 的概念，计算公式如下：

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \times 100\% \quad (2)$$

其中，TP 表示模型预测为正类实际为正类的样本数，FP 表示模型预测为正类实际为负类的样本数，FN 表示模型预测为负类实际为正类的样本数，如公式(1)、(2)所示。

$$AP = \int_0^1 P(R) dR \quad (3)$$

$$mAP = \frac{\sum_{i=1}^C AP_i}{C} \tag{4}$$

分别以不同召回率下的最高精确率作为横纵坐标得到一条 P-R 曲线，由公式(3)可知，对 P-R 曲线求积分即为平均精度值 AP。C 为目标类别数，因此由公式(4)可知 mAP 为所有类别的 AP 均值。

4.3. 实验结果及分析

在 3.3.2 模型训练参数优化的基础上训练 YOLOv7，输入图像分辨率为 640×640 ，初始学习率设为 $1e-4$ ，最低学习率设为 $1e-6$ ，迭代轮次(epoch)为 100，每次迭代的训练批次(batchsize)为 8。训练曲线如图 7 所示，从图中可以看出，在迭代 75 次之后损失值振荡减小并逐渐趋于收敛，选择收敛阶段的最优模型进行测试。

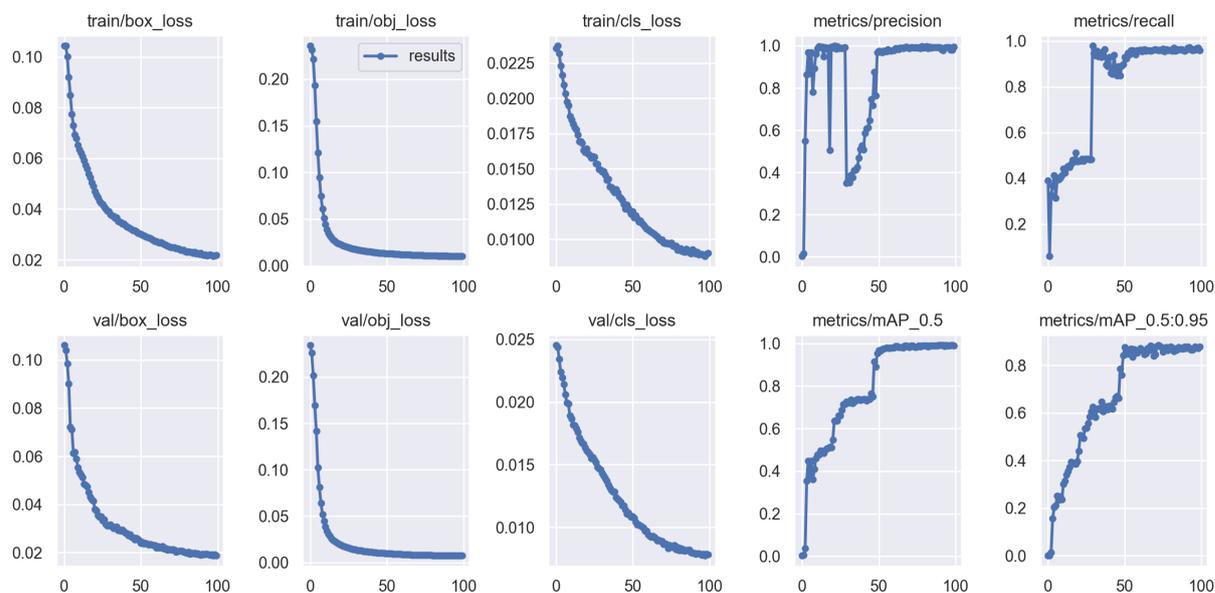


Figure 7. Algorithm training curves
图 7. 算法训练曲线

YOLOv7 在实时监控下的部分检测效果如图 8 所示，从图中可以看出，YOLOv7 算法在都能够准确的识别和定位垃圾和夹爪 2 种杂物。

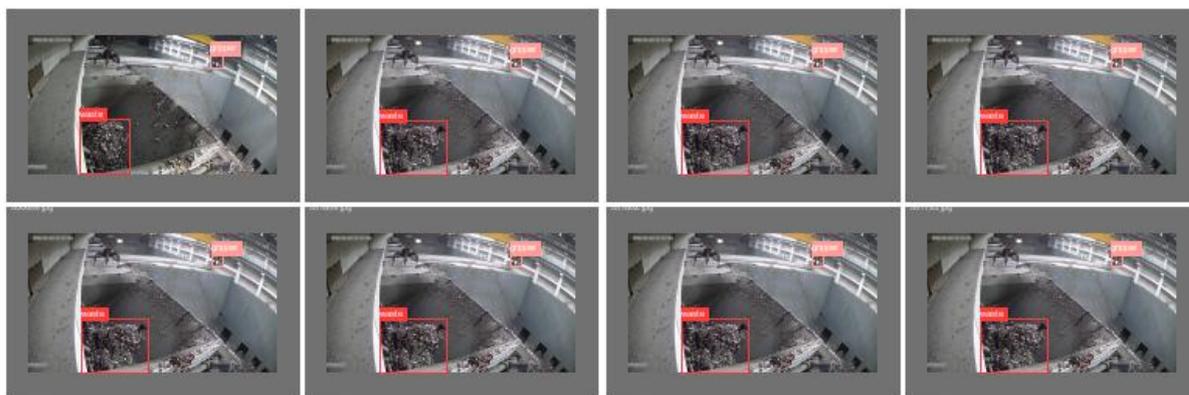


Figure 8. Test results
图 8. 检测结果

以上实验分析表明，对于实时检测垃圾和夹爪而言，YOLOv7 在检测精度方面满足检测需求，可有效检测出 2 种目标物，但其模型复杂度较高，实际推理速度在低配计算平台上未达到实时检测要求。因此，下一步需要对其进行改进，在保持高精度的同时，进一步降低模型复杂度，提高模型推理速度，达到实时检测的要求，从而更好地实现算法检测精度与检测速度的平衡。

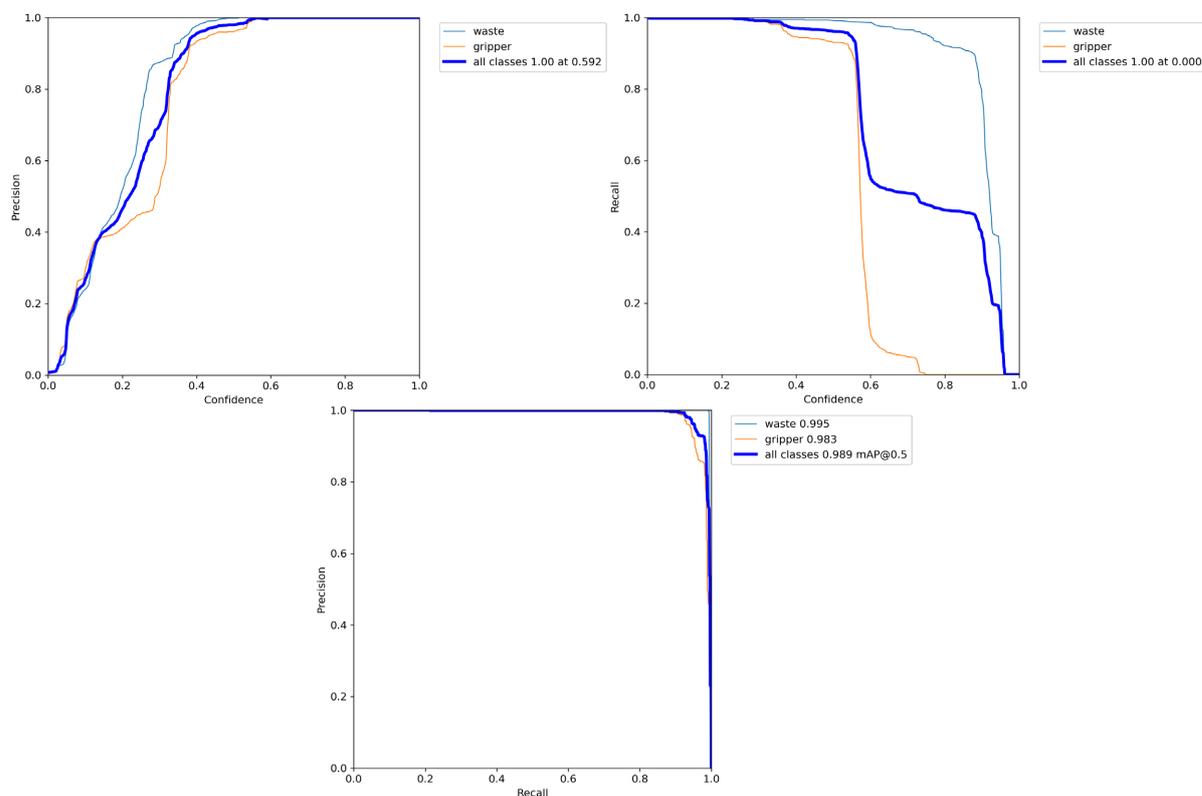


Figure 9. P-R curves

图 9. P-R 曲线

从图 9 中我们可以看出，在 YOLOv7 的检测结果中，其 Precision 曲线的值在 confidence 为 0.4~0.6 之间，逐渐趋向 1。除此之外，其 Recall 曲线的值在 0.4~0.8 之间有所下滑，最终逐渐趋向为 1。从整体上来看，它们的 PR 曲线和 Precision 曲线、Recall 曲线随置信度变化的曲线基本上一致。Precision-Recall 曲线得到 waste 类的 PR 值为 0.988，gripper 类的 PR 为 0.655，所有类别整体的 mAP0.5 的值为 0.877。

5. 结论

本文首先介绍 YOLOv7 算法的基本组成结构，阐述每一组成部分的原理。然后，构建了杂物数据集，通过 LabelImg 工具完成标签制作工作，为避免训练模型时出现过拟合或泛化性不足的问题，提出用离线数据增强方法扩充数据量，并用 K-means 算法对标签文件重聚类，得到适用于杂物数据集的先验框，有利于模型收敛，采用余弦退火算法优化学习率衰减策略，避免模型陷入局部最优。最后，在训练集上，训练 YOLOv7 模型，实验结果表明，该算法在精度方面满足要求，可以满足实时工业垃圾处理。

参考文献

- [1] Wang, C.-Y., Liao, H.-Y.M. and Yeh, I. (2022) Designing Network Design Strategies through Gradient Path Analysis. ArXiv: 2211.04800.

-
- [2] Bello, I., Fedus, W., Du, X.Z., Cubuk, E.D., Srinivas, A., Lin, T.-Y., Shlens, J. and Zoph, B. (2021) Revisiting ResNets: Improved Training and Scaling Strategies. ArXiv: 2103.07579.
- [3] Bochkovskiy, A., Wang, C.-Y. and Liao, H.-Y.M. (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection. ArXiv: 2004.10934.
- [4] Cao, Y., Geddes, T.A., Yang, J.Y.H. and Yang, P.Y. (2020) Ensemble Deep Learning in Bioinformatics. *Nature Machine Intelligence*, **2**, 500-508. <https://doi.org/10.1038/s42256-020-0217-y>
- [5] Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A. and Zagoruyko, S. (2020) End-to-End Object Detection with Transformers. In: Vedaldi, A., Bischof, H., Brox, T. and Frahm, J.M., Eds., *Proceedings of the European Conference on Computer Vision (ECCV)*, Vol. 12346, 213-229. https://doi.org/10.1007/978-3-030-58452-8_13
- [6] Chen, K., Lin, W.Y., Li, J.G., See, J., Wang, J. and Zou, J.N. (2020) AP-Loss for Accurate One-Stage Object Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, **43**, 3782-3798. <https://doi.org/10.1109/TPAMI.2020.2991457>
- [7] Chen, Z., Duan, Y.C., Wang, W.H., He, J.J., Lu, T., Dai, J.F. and Qiao, Y. (2022) Vision Transformer Adapter for Dense Predictions. ArXiv: 2205.08534.
- [8] Choi, J., Chun, D., Kim, H. and Lee, H.-J. (2019) Gaussian YOLOv3: An Accurate and Fast Object Detector Using Localization Uncertainty for Autonomous Driving. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 502-511. <https://doi.org/10.1109/ICCV.2019.00059>
- [9] Dai, X.Y., Chen, Y.P., Xiao, B., Chen, D.D., Liu, M.C., Yuan, L. and Zhang, L. (2021) Dynamic Head: Unifying Object Detection Heads with Attentions. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 7373-7382. <https://doi.org/10.1109/CVPR46437.2021.00729>
- [10] Ding, X.H., Chen, H.H., Zhang, X.Y., Huang, K.Q., Han, J.G. and Ding, G.G. (2022) Reparameterizing Your Optimizers Rather than Architectures. ArXiv: 2205.15242.
- [11] Ding, X.H., Guo, Y.C., Ding, G.G. and Han, J.G. (2019) ACNet: Strengthening the Kernel Skeletons for Powerful CNN via Asymmetric Convolution Blocks. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 1911-1920. <https://doi.org/10.1109/ICCV.2019.00200>
- [12] Ding, X.H., Zhang, X.Y., Han, J.G. and Ding, G.G. (2021) Diverse Branch Block: Building a Convolution as an Inception-Like Unit. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 10886-10895. <https://doi.org/10.1109/CVPR46437.2021.01074>