基于 YOLOv8 和 OCR 的快递盒识别算法研究与应用

范晓菁 [1]
FAN Xiaojing

摘 要 随着电商及快递行业发展，物流分拣逐渐向智能化、无人化方向发展，如何稳定高效地定位快递盒并识别快递单号文字显得尤为重要。基于此，文章提出了一种基于 YOLOv8 和 OCR（optical character recognition）的快递盒识别算法，首先采集 200 张各种快递盒图片并标定建立数据集，为使最终识别模型适应不同光照条件，通过对色温和亮度对基础数据图片集进行扩充，然后用 YOLOv8 进行训练得到最优模型，并验证不同色温和亮度条件下检测精度都大于 95%。此基础上采用 PaddleOCR 完成文字提取和分类，提取所需要的目标快递盒信息，并在中国计算机设计大赛人工智能挑战赛智慧物流专项赛比赛平台 LEO 智能移动抓取机器人上部署了以上算法，通过实际竞赛验证了本文算法的有效性。

关键词 YOLOv8；OCR；PaddleOCR；PaddlePaddle；快递盒识别

doi：10.3969/j.issn.1672-9528.2025.06.011

0 引言

随着电商和物流行业的快速发展，快递盒的自动识别与分拣需求日益增加。快递盒识别算法通过计算机视觉和深度学习技术，能够高效、准确地识别快递盒的形状、尺寸、条码等信息，从而提升物流效率。目前常用的识别方法大多基于颜色和形状的特征提取 [1-2]，深度学习建立卷积神经网络 [3]。基于颜色和形状的提取方法虽然可以对快递盒进行初步定位，但在较复杂的外界情况下，容易受到较严重干扰。基于深度学习的卷积神经网络具有较强的鲁棒性，需要大量数据和计算资源支持。

对于文字识别，通常使用传统 Open-VC 软件包进行模板匹配或者使用训练好的模型进行光学字符识别 [4]（optical character recognition, OCR）基于模板匹配的方法尽管拥有较好的通用性，但在面对存在不同大小、不同排列或者倾斜角度时会存在误判。而基于 OCR 的方法可以自动获取文字信息。

YOLOv8 是 YOLO（you only look once）系列算法的最新稳定版本，因其在实时处理速度和高精度上的卓越表现而备受关注 [5]。YOLOv8 在其前系列的基础上进行了多项改进，使其在速度、准确性和效率方面都有显著提升。YOLOv8 的整体主要架构如图 1 所示，包括 3 个主要部分：主干网络（Backbone）、特征增强网络（Neck）和检测头（Detection Head）。每个部分在目标检测过程中发挥着重要作用。

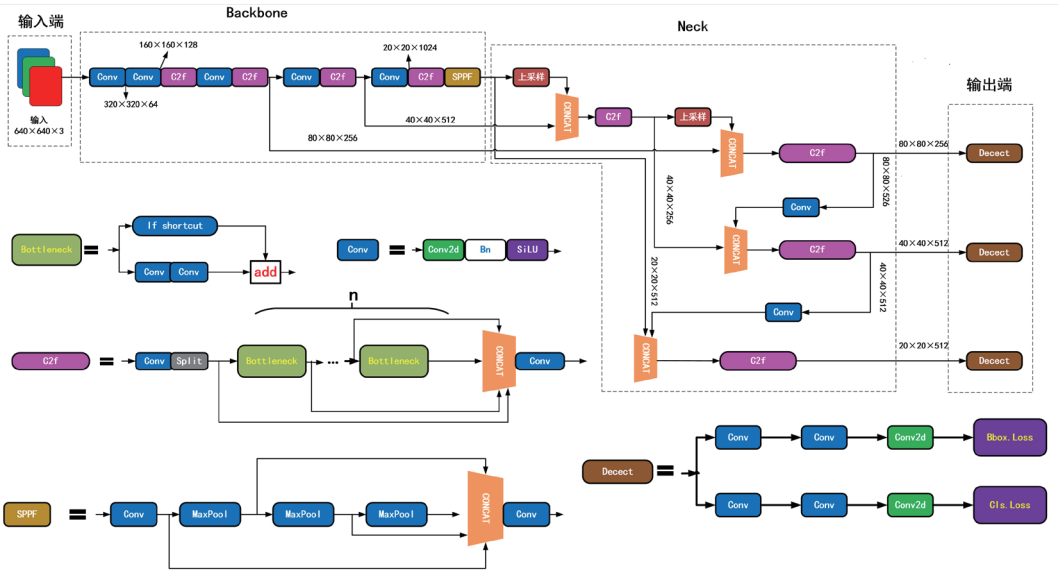

图 1 YOLOv8 的整体架构

1. 常州旅游商贸高等职业技术学校 江苏常州 213032

主干网络主要用于特征提取，通过一系列的卷积模块（Conv）、C2f 模块和 SPPF 模块完成特征提取。其中，YOLOv8 引入新的 C2f 模块用于特征提取；特征增强网络（Neck）：YOLOv8 主要采用 PA-FPN 的思想，但去掉 PA-FPN 上采样阶段的一些卷积操作；检测头：YOLOv8 引入了一个新的模块，称为解耦头（Decoupled Head），将分类分支与定位分支分离，缓解了分类和回归任务之间的固有冲突。

YOLOv8 因其优越性能成为目标检测任务中的重要工具，应用场景广泛，在交通领域的车辆检测 [6-7] 和交通标志识别 [8]、工业领域的芯片等各类产品缺陷检测 [9-12]、农业领域的农作物种类识别 [13-14]、井下人员检测 [15]、复杂场景下人体行为识别 [16] 和回收垃圾目标识别 [17] 等场景都能显著提高检测性能。

针对快递盒识别问题，本文提出了基于 YOLOv8 和 OCR 的快递盒识别算法。在标定好快递盒数据集后，为使最终识别模型适应不同光照条件，将基础数据图片集通过改变色温和亮度进行扩充，然后 YOLOv8 进行快递盒定位识别训练，在取得精度较高的模型后，采用百度 PaddleOCR 对快递单号进行识别，提取文字中的重点信息。百度推出的 PaddleOCR 是一个基于 PaddlePaddle[18] 深度学习框架的专业且开源的 OCR 工具包，能够实现文字检测、文字识别、关键点检测等各种 OCR 任务。最终在中国计算机设计大赛人工智能挑战赛智慧物流专项赛比赛平台 LEO 智能移动抓取机器人上部署了以上算法，在实际竞赛中获得国赛二等奖、三等奖，验证了本文算法的有效性。

1 算法详细介绍

1.1 数据集建立及图片色温处理算法

根据中国计算机设计大赛人工智能挑战赛智慧物流专项赛比赛场景要求，采用比赛用快递盒，在不同背景下，拍照采集了 200 张样本图片。

实际比赛时，光线对视觉识别影响非常大。而比赛场馆和比赛时间无法实现预测。因此快递盒识别需要考虑光照影响。为使最终识别模型适应不同光照条件，将基础数据图片集扩充到 1 000 K ~ 20 000 K 色温，从而得到 4 200 张样本图片，样本图片如图 2 所示。

图 2 图片数据集

用专用工具进行数据标定，创建 Annotations 和 ImageSets 文件夹，使用 makeTxt.py 进行数据划分，voc_label.py 生成类别标签，建立 yaml 配置文件，进行模型训练和测试。

1.2 评价指标

训练评估指标包括精确度（P）、召回率（R）、平均精度（AP）和平均精度均值（mAP）。

精确度 P 表示模型预测的正样本数量与所有检测到的样本数量之比；召回率 R 表示模型正确预测的正样本数量与实际正样本总数的比值，其计算公式为 [6]：

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{1}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{2}$$

式中：TP 代表正确预测的边界框数量；FP 代表被错误判断为正样本的数量；FN 代表未检测到的对象数量。

AP 指的是模型的平均准确率，而 mAP 则表示 AP 的平均值。变量 k 表示类别的数量。AP、mAP 计算公式为 [6]：

$$\mathrm{AP} = \int_0^1 P(R)\mathrm{d}R \tag{3}$$

$$\mathrm{mAP} = \frac{1}{k}\sum_1^k \mathrm{AP}_k \tag{4}$$

边界框回归是目标检测任务的核心环节。交并比（intersection over union, IoU）[19] 是边界框的标准评估指标，其公式为：

$$\mathrm{IoU} = \frac{\mathrm{Intersection}\left(B_{\mathrm{pred}}, B_{\mathrm{gt}}\right)}{\mathrm{Union}\left(B_{\mathrm{pred}}, B_{\mathrm{gt}}\right)} = \frac{\left|B_{\mathrm{pred}} \cap B_{\mathrm{gt}}\right|}{\left|B_{\mathrm{pred}} \cup B_{\mathrm{gt}}\right|} \tag{5}$$

式中：$B_{\mathrm{pred}} \cap B_{\mathrm{gt}}$ 表示预测框与真实框的交集面积；$B_{\mathrm{pred}} \cup B_{\mathrm{gt}}$ 表示预测框与真实框的并集面积。

为了解决非重叠框的梯度消失问题，广义交并比（generalized iou, GIoU）在 IoU 的基础上引入了最小闭合区域 C[19]，其公式为：

$$\mathrm{GIoU} = \mathrm{IoU} - \frac{\left|C - \left(B_{\mathrm{pred}} \cap B_{\mathrm{gt}}\right)\right|}{|C|} \tag{6}$$

式中：C 为包含 B_{pred} 和 B_{gt} 的最小矩形区域面积。

距离交并比（distance intersection over union, DIoU）在 IoU 的基础上增加了中心点距离的惩罚项，可以让边界框更快地收敛 [20]，其公式为：

$$\mathrm{DIoU} = \mathrm{IoU} - \frac{\rho^2\left(b_{\mathrm{pred}}, b_{\mathrm{gt}}\right)}{c^2} \tag{7}$$

式中：$\rho^2\left(b_{\mathrm{pred}}, b_{\mathrm{gt}}\right)$为预测框与真实框中心点的欧氏距离；$c$ 为最小闭合区域 C 的对角线长度。

完全交并比（complete iou,CIoU）在 DIoU 基础上，增加宽高比相似性约束，同时优化中心点、重叠面积和宽高比，提高回归精度 [20]。CIoU 公式为：

$$\mathrm{CIoU} = \mathrm{DIoU} - \alpha v \tag{8}$$

式中：$v=\frac{4}{\pi^2}(\arctan\frac{w_{gt}}{h_{gt}}-\arctan\frac{w_{pred}}{h_{pred}})^2$；$\alpha=\frac{v}{1-IoU+v}$；$w$、$h$ 为框的宽度和高度；v 为宽高比差异的惩罚项（通过 arctan 转化为角度差异）。

IoU 损失和广义 IoU（GIoU）损失虽有利于 IoU 指标，但仍存在收敛速度慢和回归不精准的问题。距离 IoU（DIoU）损失，通过引入预测框与目标框之间的归一化距离，其训练收敛速度显著快于 IoU 和 GIoU 损失。完全 IoU（CIoU）损失综合了边界框回归中的 3 个几何因素（重叠面积、中心点距离和宽高比），从而实现更快收敛和更优性能。因此本文选择 CIoU 损失评估边界框。

1.3 文字识别

在识别到快递盒后，需要进行文字识别。本文采用 PaddleOCR 进行 OCR 检测得到对应的结果，其中在 PaddleOCR 中使用文本检测、方向分类和文字识别串联推理由此得到最佳化识别效果。实际应用时考虑光线等各种干扰，需要用 OpenVC 进行预处理来提高识别率和稳定性。具体处理步骤如下：

（1）首先将图像进行灰度处理，变成灰色图像。然后降噪处理，使每个像素点邻域内像素值加权平均，呈现出平滑图像。最后二值化，凸显字迹轮廓。

（2）修复因照片模糊导致的字迹不清，需要进行对二值化的图像先膨胀操作，以连接断裂的字迹。接着进行腐蚀操作，消除膨大的字迹。

（3）自适应阈值处理。划分图像为小区域，为每个小区域计算阈值，加大文字和背景区分。

（4）选择快递盒子特定区域，采用 PaddleOCR 进行文字识别，排除选择区域外的文字。选择区域内识别到的文字轮廓，提取、调整编码格式，最后输出结果。

2 实验与应用

因数据集较少，为降低训练环境硬件要求，尝试用普通笔记本电脑进行训练，使学生也具备进行机器学习训练的实验环境。采用一台 Think T480 笔记本进行训练，实验环境如表 1 所示。

表 1 实验环境

环境项	环境规格
操作系统	Win10
编程语言	Python3.9
深度学习框架	PyTorch 2.5.1
CPU	Intel Core(TM) i7-8550U 1.80 GHz
内存	24 GB

训练时，图片大小为 640 px×640 px，训练轮次为 500 轮，批量大小为 4，使用 CIoU 损失评估边界框，网络初始学习率为 0.001。训练损失函数和精度指标数据如图 3 所示。

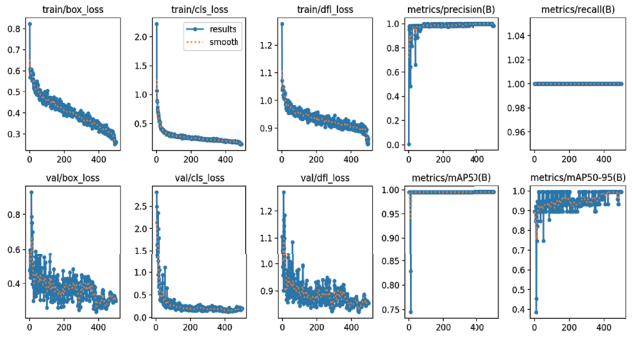

图 3 训练损失函数和精度指标图

得到最优模型后，基于 PaddleOCR 采用上文步骤进行文字识别，识别效果如图 4 所示。但是在进行全面测试时发现，光线亮度对识别效果也有影响。

图 4 不同色温图片识别文字效果

为此，再将数据集图片根据亮度系数进行扩充，再用 YOLOv8 进行训练。重新训练后同一图片不同亮度情况下也能进行快递盒识别和文字识别，效果如图 5 所示。

图 5 不同亮度图片识别文字效果

训练得到最优模型后，将模型部署到中国计算机设计大赛人工智能挑战赛智慧物流专项赛比赛平台 LEO 智能移动抓取机器人上，实现了各种环境的快递盒检测和文字识别，如图 6 所示。在实际比赛中获得 2023 年（第 16 届）中国大学生计算机设计大赛二等奖、2024 年（第 17 届）中国大学生计算机设计大赛三等奖。

图 6 LEO 机器人识别并抓取快递盒

3 结论

针对快递盒识别问题，本文提出了环境适应度更强的基于 YOLOv8 和 OCR 的快递盒识别算法。在标定好快递盒数据集后，为了使最终识别模型适应不同环境，将基础数据图片集通过改变色温和亮度进行扩充，然后 YOLOv8 进行快递盒定位识别训练，再采用图形预处理和 PaddleOCR 工具对快递单号进行识别，提取文字中的重点信息。最终在中国计算机设计大赛人工智能挑战赛智慧物流专项赛比赛平台 LEO 智能移动抓取机器人上部署了以上算法，在实际竞赛中验证了本文算法的有效性。环境对图像的影响包括亮度、白平衡（即色温）、对比度、饱和度、高光、阴影等，针对一张基础图像，可以从上述几个方面扩充图片，使得数据集更加完备，这样训练后的模型适应性更强。

参考文献：

[1] 胡文 , 马玲玉 . 基于 OpenCV 手机拍照快递单文字识别的研究 [J]. 哈尔滨商业大学学报 (自然科学版), 2015, 31(5): 564-568.

[2] 孙川 , 云利军 , 邓维轩 . 基于结构特征的快递单手机号码识别算法研究 [J]. 云南师范大学学报 (自然科学版), 2017, 37(1): 37-42.

[3] 郭广振 , 张丰收 . 基于机器视觉的快递面单识别 [J]. 包装与食品机械 , 2023,41(3):65-69.

[4] 王珂 , 杨芳 , 姜杉 . 光学字符识别综述 [J]. 计算机应用研究 , 2020, 37(S2):22-24.

[5] SOLAWETZ J, FRANCESCO. What is YOLOv8? a complete guide[EB/OL].(2024-10-23)[2024-12-26].https://blog.roboflow.com/what-is-yolov8/.

[6] WANG H, LIU C Y, CAI Y F, et al. YOLOv8-QSD: an improved small object detection algorithm for autonomous vehicles based on YOLOv8[J/OL]. IEEE transactions on instrumentation and measurement, 2024[2024-11-22].https://ieeexplore.ieee.org/document/10474434.DOI:10.1109/TIM.2024.3379090.

[7] 马明扬 , 宋淑彩 , 赵一航 , 等 . 基于 YOLOv8 的车辆类型识别研究 [J]. 河北建筑工程学院学报 , 2024, 42(1): 248-251.

[8] DING L. SlimEMA-YOLOv8: enhanced traffic sign recognition for autonomous driving using EMA and slim-neck in YOLOv8[C]//2024 IEEE 6th International Conference on Internet of Things, Automation and Artificial Intelligence (IoTAAI). Piscataway: IEEE, 2024:26-28.

[9] 张恒 , 程成 , 袁彪 , 等 . 基于 YOLOv8-EA-FPNs 的芯片缺陷检测方法研究 [J]. 电子测量与仪器学报 , 2023, 37(5): 36-45.

[10] HU Y X, HONG Y, WANG S T,et al. A YOLOv8s model integrating GELAN module and HAT attention for PCB defect detection[C/OL]// 2024 6th International Conference on Internet of Things, Automation and Artificial Intelligence (IoTAAI).Piscataway:IEEE,2024[2024-06-19].https://ieeexplore.ieee.org/document/10675021.DOI:10.1109/IoTAAI62601.2024.10692510.

[11] 王震洲 , 杨榕 , 宿景芳 . 改进 YOLOv8 的烟盒缺陷检测算法 [J]. 电子测量技术 , 2024, 47(13): 110-119.

[12] 周秋航 . 基于深度学习的烟支表面缺陷检测算法研究 [J]. 信息技术与信息化 ,2024(10):209-213.

[13] GAI R L, LIU Y, XU G H. TL-YOLOv8: a blueberry fruit detection algorithm based on improved YOLOv8 and transfer learning[J]. IEEE access, 2024,12:86378 - 86390.

[14] NIU S Q, XU X L, LIANG A, et al. Research on a lightweight method for maize seed quality detection based on improved YOLOv8[J]. IEEE access, 2024,12:32927-32937.

[15] 倪云峰 , 霍洁 , 侯颖 , 等 . 基于 YOLOv8-OCR 的井下人员检测算法 [J]. 无线电工程 , 2024, 54(8): 1847-1853.

[16] KHEKAN A R, AGHDASI H S, SALEHPOUR P. Fast and high-precision human fall detection using improved YOLOv8 model[J]. IEEE access, 2024,13:5271 - 5283.

[17] 鲁振伟 , 尚雅层 . 改进 YOLOv8 算法的可回收垃圾目标检测研究 [J]. 信息技术与信息化 ,2024(6):56-59.

[18] BI R, XU T T, XU M X, et al. PaddlePaddle: a production-oriented deep learning platform facilitating the competency of enterprises[C]//2022 IEEE 24th Int Conf on High Performance Computing & Communications. Piscataway:IEEE, 2022:18-20.

[19] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019:15-20.

[20] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]// Proceedings of the AAAI Conference on Artificial Intelligence, Palo Alto: AAAI, 2020: 12993-13000.

【作者简介】

范晓菁（1981—），女，江苏常州人，硕士，讲师，研究方向：计算机视觉、计算机应用。

（收稿日期：2025-03-02 修回日期：2025-06-09）