

文章编号: 2095-2163(2023)01-0118-05

中图分类号: TP391

文献标志码: A

# 基于轻量化 YOLO 模型的植物叶片识别

常黎玫<sup>1</sup>, 丁学文<sup>1,3</sup>, 杨旭<sup>1</sup>, 孙盼盼<sup>1</sup>, 蔡鑫楠<sup>1</sup>, 董国军<sup>2</sup>

(1 天津职业技术师范大学 电子工程学院, 天津 300222; 2 天津市高速铁路无线通信企业重点实验室, 天津 300450;

3 天津云智通科技有限公司, 天津 300350)

**摘要:** 为了提升植物叶片的检测能力,减少繁琐的人工成本,本文构建了包含 10 类常见植物叶片的数据集。应用基于 pytorch 框架的 YOLOv5 网络模型对叶片数据集进行训练,并通过训练模型对叶片测试图像进行检测,在识别速度和精度上取得了较好的效果。实验结果表明,本文检测方法对叶片识别的平均精度为 93%,识别速度较快,有效解决了传统植物识别方法中分类器耗时长、准确率低等问题。

**关键词:** 植物识别; YOLOv5; 深度学习; 神经网络; 平均精度值

## Plant leaf recognition based on lightweight YOLO model

CHANG Limei<sup>1</sup>, DING Xuewen<sup>1,3</sup>, YANG Xu<sup>1</sup>, SUN Panpan<sup>1</sup>, CAI Xinnan<sup>1</sup>, DONG Guojun<sup>2</sup>

(1 School of Electronic Engineering, Tianjin University of Technology and Education, Tianjin 300222, China;

2 Tianjin High Speed Railway Wireless Communication Enterprise Key Laboratory, Tianjin 300450, China;

3 Tianjinyunzhitong Technology Co., Ltd, Tianjin 300350, China)

**[Abstract]** In order to improve the detection capability of plant leaves and reduce the tedious labour cost, a dataset containing 10 types of common plant leaves was constructed in this paper. A YOLOv5 network model based on the pytorch framework was applied to train the leaf dataset, and the leaf test images were detected by the training model, which achieved better results in terms of recognition speed and accuracy. The experimental results show that the average accuracy of the detection method in this paper for leaf recognition is 93% and the recognition speed is fast, which effectively solves the problems of time-consuming classifier and low accuracy rate in traditional plant recognition methods.

**[Key words]** Plant recognition; YOLOv5; deep learning; neural networks; mAP

## 0 引言

植物是大自然基本资源,是食物、药物的主要来源。植物作为大自然中的初级生产者,可以各种形式储存能量,供生物界的能量需求,其在维持地球上的生物和物质循环平衡中,起着不可替代的作用。如:植物为地球上的生物提供栖息地和生活场所;植物在水土保持和净化生物圈的无机环境中起着不可或缺的作用;植物对于人类生产生活有着重要的经济地位,主要表现在药用和食品方面,是人类生存的基本而又无法替代的部分。作为农业大国,提高植物业和农业生产需要大量植物的精细数据。因此,植物分类识别对于人类生产生活具有重要意义和作用。

随着 20 世纪计算机与人工智能技术的兴起与

发展,基于深度学习的目标检测算法已逐渐代替传统的识别方法。深度学习的目标检测算法分为一阶段(one-stage)目标检测算法和二阶段(two-stage)目标检测算法。二阶段目标检测算法以 R-CNN<sup>[1]</sup>系列为代表,先使用卷积神经网络(Convolutional Neural Networks, CNN)从图像中提取出一些候选框,然后再对这些候选框进行逐一甄别,筛选出置信度高的候选框。该算法相当于进行了两次筛选,因此检测的准确率较高、计算量低,但检测的速度较慢。一阶段目标检测算法以 YOLO 系列<sup>[2-5]</sup>和 SSD<sup>[6]</sup>为代表,这种方法直接将图像数据送入网络,提取一次特征,然后使用卷积神经网络将特征直接提取候选框,整个过程只要一步就可完成。因此,单阶段目标检测的优势是检测速度快。

**基金项目:** 天津市科委科技特派员项目(20YDTPJC01110)。

**作者简介:** 常黎玫(1996-),女,硕士研究生,主要研究方向:机器视觉。

**通讯作者:** 丁学文 Email: dingxw1@126.com

**收稿日期:** 2022-04-26

近年来,国内外研究者基于图像处理和机器学习的植物识别进行了多方位的研究。如:文献[7]中提出,将叶片形状特征和纹理特征相结合,共提取了 2 182 维特征,再使用线性评判分析与主成分分析相结合的方式,将叶片进行特征降维,降维后的训练集取得了不错的识别结果。文献[8]构建了一种基于残差连接的 Alexnet 卷积神经网络,对于多种植物叶片识别不准、计算效率低等问题得到了明显的改善和提升。文献[9]在 P1antCLEF 2016 比赛中发现使用微调的 AlexNet、VGGNet 和 GoogLeNet 模型来识别植物种类的方法优于其它方法。文献[10]将 LBP、SURF 和 HOG 3 种纹理特征单独输入到支持向量机分类器中对自建的草本植物数据集进行分类识别,结果表明 HOG 和 LBP 的识别性能优于 SURF<sup>[10]</sup>。

目前,叶片识别使用的很多模型都需消耗大量的计算资源,只能部署在服务器端,导致其普及性和实时性较差,所以对于叶片识别具有一定的局限性。而 YOLO 神经网络模型由于其算法的改进性和迁移性强,在快速识别植物叶片和移动端部署上有着更好的表现和应用前景。本文利用基于 pytorch 框架的 YOLOv5 中 v5s 模型来训练植物叶片数据,以满足植物叶片识别的准确性和高效性。

### 1 神经网络模型概述

YOLOv5 是 YOLO 系列不断创新的一阶段目标检测网络。在 YOLOv5 模型的代码中,4 个版本的检测网络分别为:YOLOv5x、YOLOv5l、YOLOv5m、YOLOv5s,可以满足不同量级的设备使用。其中,YOLOv5s 是该系列中深度最小,并且特征图的宽度也是最小的网络。YOLOv5 沿用了 YOLOv4 的部分

方法。例如:利用 Mosaic 方法进行数据增强;针对不同数据集计算不同的初始先验框(anchor)长宽比;在测试时,对原始图片添加最少的边框,以提高推理速度;改进损失函数 Generalized Intersection over Union (GIoU) Loss 和预测框筛选的 Distance-IoU(DIOU)\_NMS 等等。

#### 1.1 YOLOv5 模型结构

YOLOv5 网络结构由输入端(Input)、Backbone、Neck、Prediction 组件组成,各部分完成的主要功能所述如下:

(1) Input: YOLOv5 沿用 YOLOv4 中 Mosaic 数据增强方式来丰富数据集,降低硬件需求,减少 GPU 的使用。自适应锚框计算功能被嵌入到整个训练代码中,可自行调节开关,自适应图片缩放,提高了推理阶段目标检测的速度。

(2) Backbone: Focus 模块是 YOLOv5 特有结构,其中的关键是将特征图进行分块切片操作,之后经过多层卷积池化、跨阶段局部网络(Cross Stage Partial Network, CSPNet)<sup>[11]</sup>和特征金字塔池化(Spatial Pyramid Pooling, SPP)<sup>[12]</sup>结构,将图像特征传递到下一组块。

(3) Neck:沿用了 YOLOv4 的特征金字塔网络(Feature Pyramid Networks, FPN)<sup>[13]</sup>和路径聚合网络(Perceptual Adversarial Network, PAN)<sup>[14]</sup>结构。但不同的是,在 YOLOv5 的 Neck 结构中,采用了借鉴 CSPnet 设计的 CSP2 结构,加强了信息传播,具有准确保留空间信息能力。

(4) Prediction:改进损失函数 GIoU-Loss,可以将一些遮挡重叠的目标准确识别出来,进而能够解决传统 IOU 存在的缺陷。

YOLOv5s 网络结构如图 1 所示。

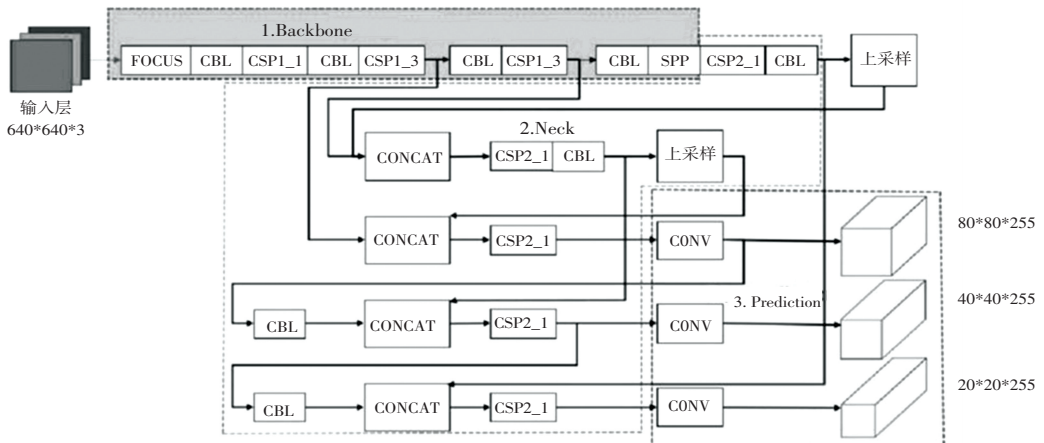


图 1 YOLOv5s 网络结构图  
Fig. 1 YOLOv5s Network Structure Diagram

## 1.2 YOLOv5 损失函数

在 YOLOv5 网络中使用  $IoU$  系列损失函数中的  $CIoU$  loss 来计算目标框和预测框的边界框损失 (Box)。其中标框和预测框的置信度损失 (Objectness) 与分类损失 (Classification) 则是使用二元交叉熵损失函数 (Binary Cross Entropy loss, BCE loss) 来计算。

(1) 边界框损失:

$$CIoU = IoU - \rho^2/c^2 - \alpha v \quad (1)$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w^p}{h^p} \right)^2 \quad (2)$$

$$loss_{CIoU} = 1 - CIoU \quad (3)$$

其中,  $IoU$  是预测框和目标框两个方框重叠面积与相并部分面积的比值;  $\rho$  为预测框和目标框的

中心点距离;  $v$  用来表示长宽比的相似度;  $\alpha$  为  $v$  的影响系数。

(2) 置信度损失与分类损失:

$$Loss = - \sum_{i=1}^N (y_i \log(p(x_i)) + (1 - y_i) \log(1 - p(x_i))) \quad (4)$$

其中,  $y$  为真实标签,  $p(x)$  为模型输出。

## 2 实验

### 2.1 数据集

实验选用 10 种植物类别的叶片作为数据集, 将实验选用的样本图片按照 3:1 比例划分训练集和测试集, 其中部分植物叶片数据集如图 2 所示。样本图片尺寸均为 416×416 分辨率, 图像背景纯白无阴影。

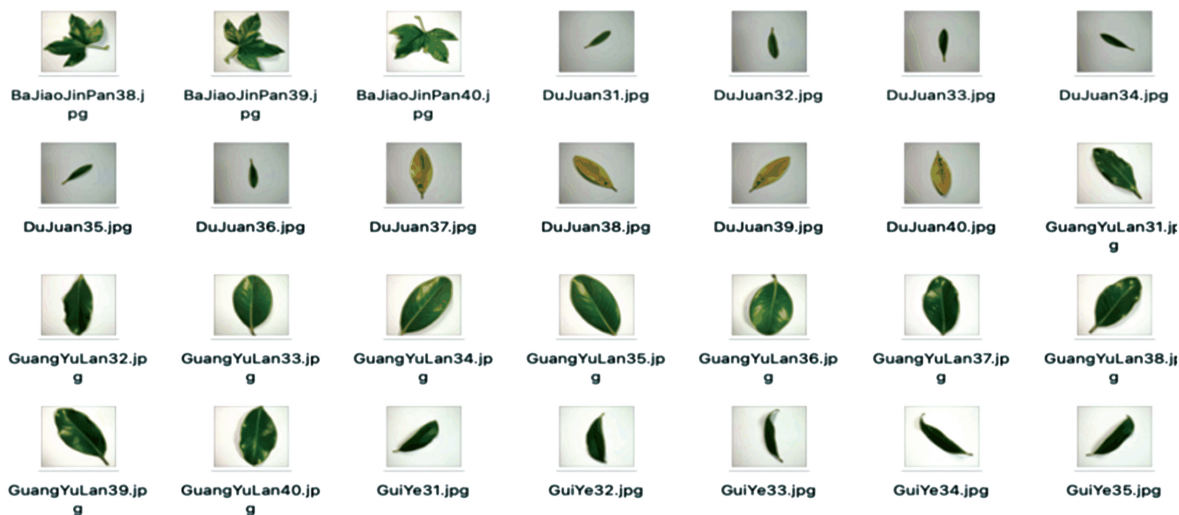


图 2 部分植物叶片数据集

Fig. 2 Partial plant leaf dataset

由于训练数据集数量较小, 网络中使用了 Mosaic 数据增强。随机从叶片数据集中找出 4 张图片, 通过随机缩放、随机裁剪、随机分布的方式进行拼接, 可以丰富叶片背景, 也增加了很多小目标, 如图 3 所示。送入网络中训练, 使得新图片中可识别信息更加丰富, 同时增强了网络的鲁棒性。



(a) 随机 4 张叶片图片 (b) 随机拼接后样本图片

图 3 Mosaic 数据增强后叶片数据集

Fig. 3 Mosaic data enhanced with leaf dataset

### 2.2 实验环境与参数配置

本实验软件环境配置见表 1, 训练网络为 YOLOv5s 网络。批尺寸 (batch size) 设置为 8, 迭代次数为 300, 初始学习率为 0.01。

表 1 实验环境参数

Tab. 1 Experimental environmental parameters

实验环境	配置参数
操作系统	Windows11 64 位
GPU	NvidiaGeforceRTX3060
深度学习框架	Pytorch1.8.1
编程语言	Python3.7.0
CPU	Intel® coreTMi5

### 2.3 模型评估指标

训练时记录下边界框损失 (Box)、置信度损失 (Objectness) 和分类损失 (Classification), 可以检查是否存在过拟合现象; 也可以通过对方框、目标检测以及分类效果进行观测, 来设置合适的参数阈值。

上述 3 种损失均值越小, 训练越精准, 识别越准确, 训练效果越好。

评价一个模型性能的好坏在于评估指标的真实性, 本文选取精准率 (*Precision*)、召回率 (*Recall*), 以及目标检测算法中衡量算法效果的指标全类平均准确率 (Mean Average Precision, *mAP*) 作为评估指标。其中, *Precision* 和 *Recall* 可以清晰反映出网络对识别对象的错误识别和没有识别到的情况。其计算公式如下:

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

式中, *TP* 是样本的真实类和预测结果一样是正例数量; *FP* 是样本的真实类别是负例, 但预测结果是正例的数量; *FN* 是样本的真实类别是正例, 但预测结果是负例的数量。

全类平均精度 *mAP*, 是将所有类别检测的平均正确率 (*AP*) 进行综合加权平均而得到, 其值也可用 *Precision* 和 *Recall* 作为两轴作图后所围的面积来表示。

### 2.4 实验结果分析

使用 YOLOv5s 网络的训练损失率和验证损失率都较为平滑, 模型拟合效果较好。如图 4 所示。

训练后模型的 *Precision*、*Recall*、*mAP@0.5* 的结果如图 5 所示。

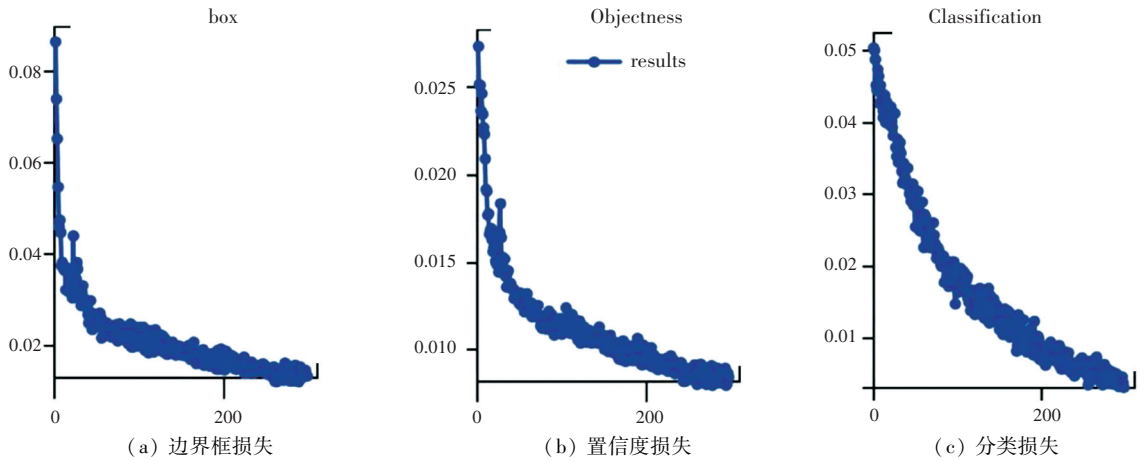


图 4 叶片训练集损失

Fig. 4 Leaf training set loss

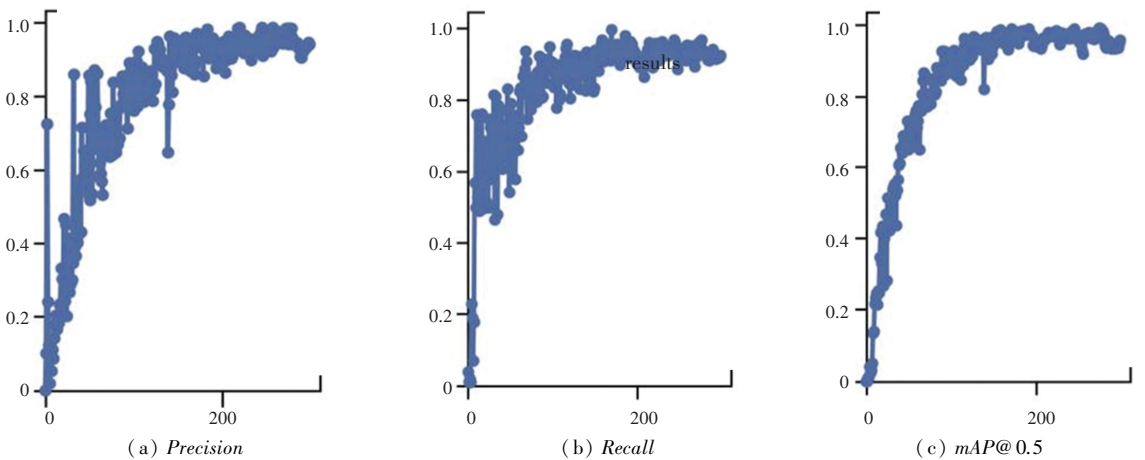


图 5 模型性能评估指标

Fig. 5 Model Performance Evaluation Metrics

由图中指标可以看出, *Precision* 和 *Recall* 在 *epoch*200 次以后波动幅度不大, *mAP@0.5* 基本稳定在 93%, 说明训练效果较好。

用本实验训练集进行识别测试, 部分结果如图 6 所示。对于目标叶片类别能够准确的进行分类, 且有较高的识别率。

基于 YOLOv5 网络训练后生成的 best 模型, 使





图6 YOLOv5 部分识别结果

Fig. 6 Partial recognition results of YOLOv5

### 3 结束语

本论文针对植物叶片识别检测任务,构建10种植物类型数据集,通过将原始叶片图像数据集随机缩放、随机裁剪、随机分布进行拼接,再进行镜像反转操作,来增强数据集并提高网络鲁棒性;使用基于Pytorch框架的YOLOv5网络模型,对叶片识别精度有较高提升,与传统人工植物叶片识别方法相比工作量大大缩减,叶片识别效率和准确率较高,网络模型可部署性强,可以满足植物学相关人员实时识别植物的要求,并减少在叶片识别中所需要的各方面成本。

### 参考文献

- [1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2014: 580-587.
- [2] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.
- [3] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 7263-7271.
- [4] REDMON J, FARHADI A. Yolov3: An incremental improvement [J]. arXiv preprint arXiv: 1804.02767, 2018.
- [5] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4:

Optimal speed and accuracy of object detection [J]. arXiv preprint arXiv:2004.10934, 2020.

- [6] LIU W, ANGUELOV D, et al. SSD: Single shot multibox detector [C]//European conference on computer vision. Springer, Cham, 2016: 21-37.
- [7] 郑一力,张露. 基于迁移学习的卷积神经网络植物叶片图像识别方法 [J]. 农业机械学报, 2018, 49(S1): 354-359.
- [8] 孙颖异,李健,时天,等. 基于改进的 AlexNet 卷积神经网络的植物叶片识别 [J]. 种子, 2020, 39(2): 77-81.
- [9] JEON W, RHEE S. Plant Leaf Recognition Using a Convolution Neural Network [J]. International Journal of Fuzzy Logic and Intelligent Systems, 2017, 17(1): 26-34.
- [10] IBRAHIM Z, SABRI N, MANGSHOR A. Leaf recognition using texture features for herbal plant identification [J]. Indonesian Journal of Electrical Engineering and Computer Science, 2018, 9(1): 152-156.
- [11] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN [C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020: 390-391.
- [12] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9): 1904-1916.
- [13] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117-2125.
- [14] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 8759-8768.

(上接第117页)

- [9] 刘显贵,王晖年,洪经纬,等. 网联环境下信号交叉口车速控制策略及优化 [J]. 交通运输系统工程与信息, 2021, 21(2): 82-90.
- [10] 鹿应荣,许晓彤,丁川,等. 车联网环境下信号交叉口车速控制

策略 [J]. 交通运输系统工程与信息, 2018, 18(1): 50-58, 95.

- [11] 徐丽萍,邓明君. 基于车路协同环境的信号交叉口车速引导方法 [J]. 交通信息与安全, 2021, 39(2): 78-86.