

基于少量标注样本的茶芽目标检测 YSVD-Tea 算法

郑子秋¹ 宋彦^{1,2} 陈霖¹ 张航¹ 宁井铭³

(1. 安徽农业大学工学院, 合肥 230036; 2. 安徽省智能农机装备工程研究中心, 合肥 230036;

3. 安徽农业大学茶树生物学与资源利用国家重点实验室, 合肥 230036)

摘要: 构建大规模茶芽目标检测数据集是一项耗时且繁琐的任务, 为了降低数据集构建成本, 探索少量标注样本的算法尤为必要。本文提出了 YSVD-Tea (YOLO singular value decomposition for tea bud detection) 算法, 通过将预训练模型中的基础卷积替换为 3 个连续的矩阵结构, 实现了对 YOLOX 算法结构的重构。通过维度变化和奇异值分解操作, 将预训练权重转换为与重构算法结构相对应的权重, 从而将需要进行迁移学习的权重和需要保留的权重分离开, 实现保留预训练模型先验信息的目的。在 3 种不同数量的数据集上分别进行了训练和验证。在最小数量的 1/3 数据集上, YSVD-Tea 算法相较于改进前的 YOLOX 算法, mAP 提高 20.3 个百分点。对比测试集与训练集的性能指标, YSVD-Tea 算法在测试集与训练集的 mAP 差距仅为 21.9%, 明显小于 YOLOX 的 40.6% 和 Faster R-CNN 的 55.4%。在数量最大的数据集上, YOLOX 算法精确率、召回率、F1 值、mAP 分别为 86.4%、87.0%、86.7% 和 88.3%, 相较于对比算法均最高。YSVD-Tea 在保证良好性能的同时, 能够更好地适应少量标注样本的茶芽目标检测任务。

关键词: 茶芽; 目标检测; 奇异值分解; 少量样本; 遗传算法; YOLOX

中图分类号: TP391.9

文献标识码: A

文章编号: 1000-1298(2024)08-0301-11

OSID:

YSVD-Tea Algorithm for Tea Bud Object Detection
Based on Few Annotated SamplesZHENG Ziqiu¹ SONG Yan^{1,2} CHEN Lin¹ ZHANG Hang¹ NING Jingming³

(1. School of Engineering, Anhui Agricultural University, Hefei 230036, China

2. Anhui Provincial Engineering Research Center of Intelligent Agricultural Machinery, Hefei 230036, China

3. State Key Laboratory of Tea Plant Biology and Utilization, Anhui Agricultural University, Hefei 230036, China)

Abstract: Constructing a large-scale dataset for tea bud object detection is a time-consuming and intricate task. To mitigate the cost of dataset construction, exploring algorithms with a minimal number of annotated samples is particularly necessary. The YOLO singular value decomposition for tea bud detection (YSVD-Tea) algorithm was introduced, which achieved the reconstruction of the YOLOX structure by replacing the basic convolution in the pre-trained model with three consecutive matrix structures. Through dimension transformation and singular value decomposition operations, pre-trained weights were converted into weights corresponding to the reconstructed algorithm structure, thereby separating the weights that require transfer learning from those that needed to be retained. This achieved the goal of preserving the general semantic information of the pre-trained model. Training and validation on three datasets of varying sizes were conducted. On the smallest 1/3 dataset, the YSVD-Tea algorithm showed a 20.3 percentage points improvement in mAP compared with the original YOLOX algorithm. Comparing performance metrics between the test and training sets, the mAP difference for the YSVD-Tea algorithm was only 21.9%, which was significantly lower than YOLOX's 40.6% and Faster R-CNN's 55.4%. In training with the largest complete dataset, the YOLOX algorithm achieved precision, recall, F1 score, and mAP of 86.4%, 87.0%, 86.7%, and 88.3%, respectively, surpassing the comparison algorithms. YSVD-Tea algorithm demonstrated superior suitability for the task of tea bud object detection, especially when

收稿日期: 2023-11-23 修回日期: 2024-02-26

基金项目: 国家重点研发计划项目(2021YFD1601102)和安徽省自然科学基金项目(2308085MC84)

作者简介: 郑子秋(1998—), 男, 硕士生, 主要从事智能农业技术与装备研究, E-mail: ahau_conan@stu.ahau.edu.cn

通信作者: 宋彦(1983—), 男, 副教授, 主要从事智能农业技术与装备研究, E-mail: songyan@ahau.edu.cn

confronted with a limited number of annotated samples.

Key words: tea bud; object detection; singular value decomposition; small sample size; genetic algorithm; YOLOX

0 引言

目标检测技术在农业领域有广泛应用^[1-3],特别是在茶叶目标检测任务中^[4-5]。为实现茶叶智能化采摘,学者们提出了多种基于深度学习的茶芽目标检测方法^[6-7]。GUI等^[8]通过引入 Ghost 模块成功轻量化了 YOLO v5 模型,并在 1 000 幅图像上完成了模型的训练验证。LI等^[9]采用增强的 YOLO v5 模型进行茶芽计数,在 4 260 幅图像上训练并实现了茶叶产量估计任务。CHEN等^[10]基于 YOLO v3 提出了茶芽和茶叶的识别与定位方法,在 2 000 幅图像上取得了 71.96% 的茶芽平均准确率。同时,王梦妮等^[11]、ZHANG等^[12]、CHEN等^[13]分别使用 1 190、1 000、4 015 幅图像构建数据集,用于茶叶检测任务的训练。

尽管上述工作取得了显著的成果,但在训练过程中往往需要构建大规模的数据集,其中包含数百至数千幅图像,以确保模型的泛化性能。由于茶叶目标小而密集,单幅图像的标注成本相对较高,构建大规模茶叶数据集的成本更高。在这一背景下,如何在少量标注样本情况下,维持目标检测模型的强泛化性能,成为一个值得研究的方向。

迁移学习旨在通过将在大规模源数据集上学到的知识应用到新的目标领域,从而提升模型的泛化能力和检测效果^[14-16]。因此,迁移学习被广泛运用在少样本训练任务中,例如,刘世晶等^[17]提出的小样本鱼图像识别算法、李子茂等^[18]提出的茶叶病害小样本识别算法等。此外,还有度量学习^[19]、元学习^[20]等方法被用于小样本学习。然而在众多小样本学习方法中,迁移学习因其相对较好的效果和易实现性而备受青睐。通常,迁移学习采用冻结主干或部分主干来保留在大规模源数据集上学到的知识,但这种方式一般按照特征层进行划分,缺乏一定的选择性。研究表明,通过训练模型的部分参数以提高泛化能力的方法是有效的^[20-21]。为了更好地分离这些有效特征,WU等^[22]和 SUN等^[23]采用奇异值分解的方式将特征和权重分解为多个部分相乘的形式,并验证了奇异值分解可以将源数据集上学到的语义信息与需要迁移学习的权重分离开,并经过大量实验证明这种分解方法对提升少量数据集的泛化能力是有效的。

本文采用奇异值分解的方法,通过奇异值分解

的方式分离语义信息和任务相关信息,有选择性地微调与迁移任务相关的权重。YSVD-Tea 算法采用卷积替换的方式,对 YOLOX^[24] 网络进行重构,并运用网络奇异值分解方法将预训练权重中的通用语义信息与迁移任务相关的权重尽可能分隔开。通过只训练与迁移任务相关的权重,达到增强 YSVD-Tea 算法泛化能力的目的,使其能够在小规模茶芽数据集上进行有效训练,一定程度上避免过拟合的风险。

1 材料与方法

1.1 材料

1.1.1 数据集制作

本研究的实验数据采集自茶叶种植园内,覆盖尽可能接近真实情境的茶叶数据。采集的场景包括不同光照条件、不同视角以及各种背景干扰。使用 Zed 相机进行数据采集,共计获得 190 幅图像,每幅图像分辨率为 1 920 像素 × 1 080 像素。使用 LabelImg 工具对这些数据进行了标记,并按照 PASCAL VOC 的数据集格式建立了与每幅图像相对应的 XML 标签文件。为了提升模型泛化能力,本研究对这 190 幅标记图像按照比例 9:10 将数据集划分为训练集和测试集。然后,在原始训练集的基础上进行了 1/2 和 1/3 随机采样,分别获得了两种更小规模的数据集。对这 3 个规模的训练集进行数据增强操作,包括缩放、翻转、高斯模糊和剪切,将训练集分别扩充至 360、180、120 幅,具体见表 1。随后,采用交叉验证的方法对训练集进行训练和验证,而测试集则用于评估模型的最终检测性能。

表 1 数据集样本数量

Tab.1 Number of samples in dataset

数据集类型	完整数据集	1/2 数据集	1/3 数据集
训练集	90	45	30
训练集增广	360	180	120
测试集	100	100	100

1.1.2 实验平台

实验中的训练和测试均在同一环境下部署和运行。CPU 采用英特尔 Xeon (R) E5 - 2680 v4 @ 2.40 GHz,GPU 采用 NVIDIA 公司的 TITAN Xp,其显存为 12 GB。机器运行内存为 16 GB,操作系统为 Ubuntu 18.04,Python 为 3.8 版本,Cuda 版本 11.1,深度学习框架使用 PyTorch 1.8.1。

1.2 方法

1.2.1 网络奇异值分解与微调

在少量标注样本上微调模型存在较大的过拟合风险。通过冻结主干网络预训练模型可以降低这种风险,但是完全冻结主干网络训练又会限制模型的迁移性能。因此,优化目标是在微调以适应茶芽目标检测任务的同时仍能保留通用语义信息,以降低过拟合风险并增强模型的泛化能力。本文通过网络奇异值分解方式将预训练中通用语义相关的权重与迁移任务相关的权重尽可能分隔开,只微调与迁移任务相关的权重,实现保留通用

语义信息的同时尽可能减少模型在迁移训练时的限制。

网络的奇异值分解包括两个主要部分:网络重构和参数重构。选取 CSPDarkNet 主干网络^[24]为基线网络,网络重构是通过卷积替换操作将 CSPDarkNet 主干网络中的所有卷积运算替换成卷积与矩阵乘法复合运算形式来实现的。如图 1 所示,CSPDarkNet 主干网络中的 Conv 3×3 卷积结构被直接替换为 Conv 3×3 卷积、Mul 3×3 矩阵乘法和 Conv 1×1 卷积结构。替换主干网络里的所有 3×3 卷积运算,从而完成网络重构。

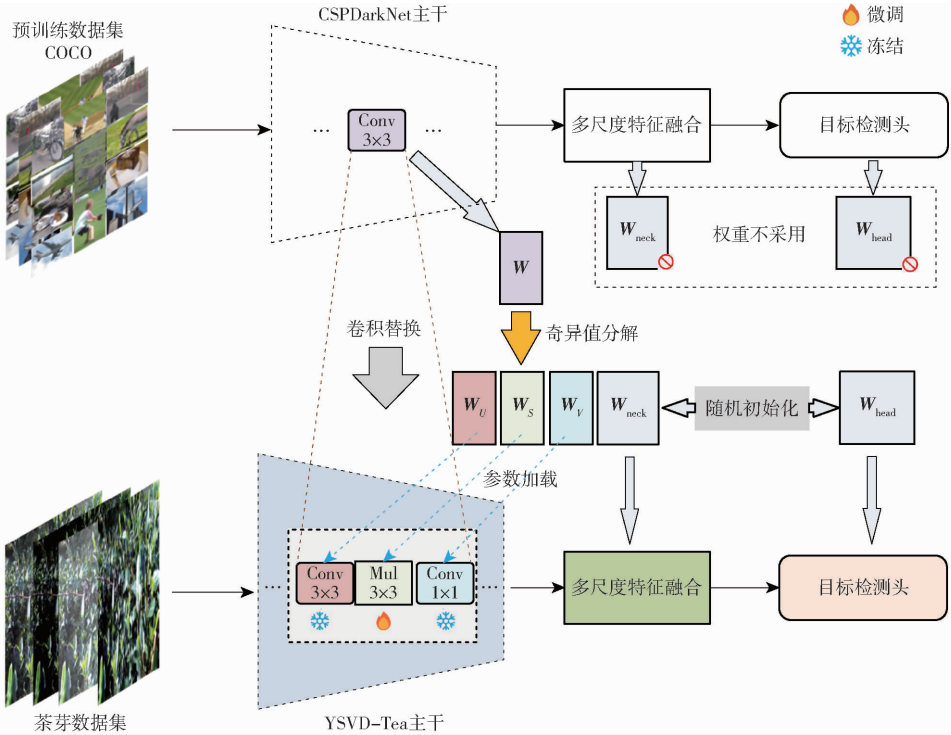


图 1 网络重构与微调

Fig.1 Network reconstruction and fine-tuning

参数重构则是将预训练权重通过奇异值分解的方式分解成不同的 3 个部分。具体而言,预训练权重是通过对其中的 Conv 3×3 卷积权重分解实现的。首先卷积权重维度调整的方法如图 2 所示,卷积层的权重张量为 $\mathbf{W} \in \mathbf{R}^{C_o \times C_i \times K \times K}$,其中 C_i 为卷积层的输入通道数, C_o 为卷积层的输出通道数, $K \times K$ 为卷积核的大小。权重张量中维度为 $C_i \times K \times K$ 的单个卷积核被展平成维度 $1 \times C_i K^2$ 的向量。因此 C_o 个卷积核被展平成权重矩阵 $\mathbf{W}' \in \mathbf{R}^{C_o \times C_i K^2}$ 。然后将重构后的权重矩阵 $\mathbf{W}'$ 进行奇异值分解,数学表达式为

$$\mathbf{W}' = \mathbf{U} \mathbf{S} \mathbf{V}^T \tag{1}$$

其中 $\mathbf{U} \in \mathbf{R}^{C_o \times R}$ $\mathbf{S} \in \mathbf{R}^{R \times R}$
 $\mathbf{V}^T \in \mathbf{R}^{R \times C_i K^2}$ $R = \min(C_o, C_i K^2)$

$\mathbf{U}$ 和 $\mathbf{V}$ 是权重矩阵 $\mathbf{W}'$ 奇异值分解后得到的西

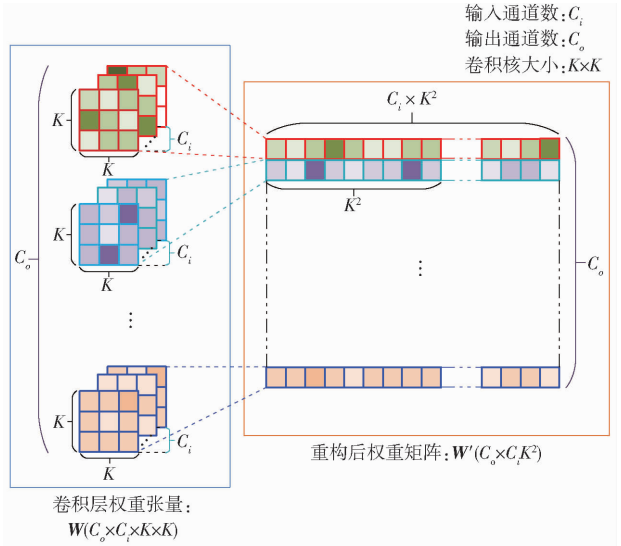


图 2 卷积权重维度调整

Fig.2 Convolutional weight dimension adjustment

矩阵, S 是对角线为奇异值的对角阵。参考图 1 中参数重构部分, 权重 W 经过参数重构后被分解为权重 W_U 、 W_S 、 W_V , 这些权重分别作为 $\text{Conv } 3 \times 3$ 卷积、 $\text{Mul } 3 \times 3$ 矩阵乘法和 $\text{Conv } 1 \times 1$ 卷积权重的初始值。多尺度融合和检测头的权重参数通常与迁移任务的关联性较强, 所以不采用预训练模型的权重, 而使用随机初始化的方式设定初始权重。 U 子空间和 V 子空间分别表示丰富的语义先验信息, 而 S 空间则涵盖了这些语义先验的综合特性^[22-23]。因此考虑冻结 U 子空间和 V 子空间权重, 只对 S 空间权重进行微调操作。从而实现保留预训练模型中的语义先验信息, 提高算法泛化能力的目的。

1.2.2 YSVD-Tea 目标检测算法

使用网络奇异值分解方法构建 YSVD-Tea 算法。该算法主要包括预处理、主干网络、多尺度特征融合模块和检测头。具体的结构如图 3 所示, 预处理部分对输入图像进行了灰边填充、Mosaic 数据增强和尺寸变换操作, 将其由 $1\,920 \times 1\,080 \times 3$ 维度转换为 $1\,280 \times 1\,280 \times 3$ 维度, 然后输入主干网络。本文提出了 Conv-SVD 结构, 该结构是 YSVD-Tea 主干网络中的基本卷积块, 它是由卷积 Conv-U 、矩阵乘法 Mul-S 和卷积 Conv-V 组成。此外, 本文还引入了 CBS-SVD 块, 它是由 Conv-SVD 结构、批归一化层 BN 和 SiLU 激活函数组成。

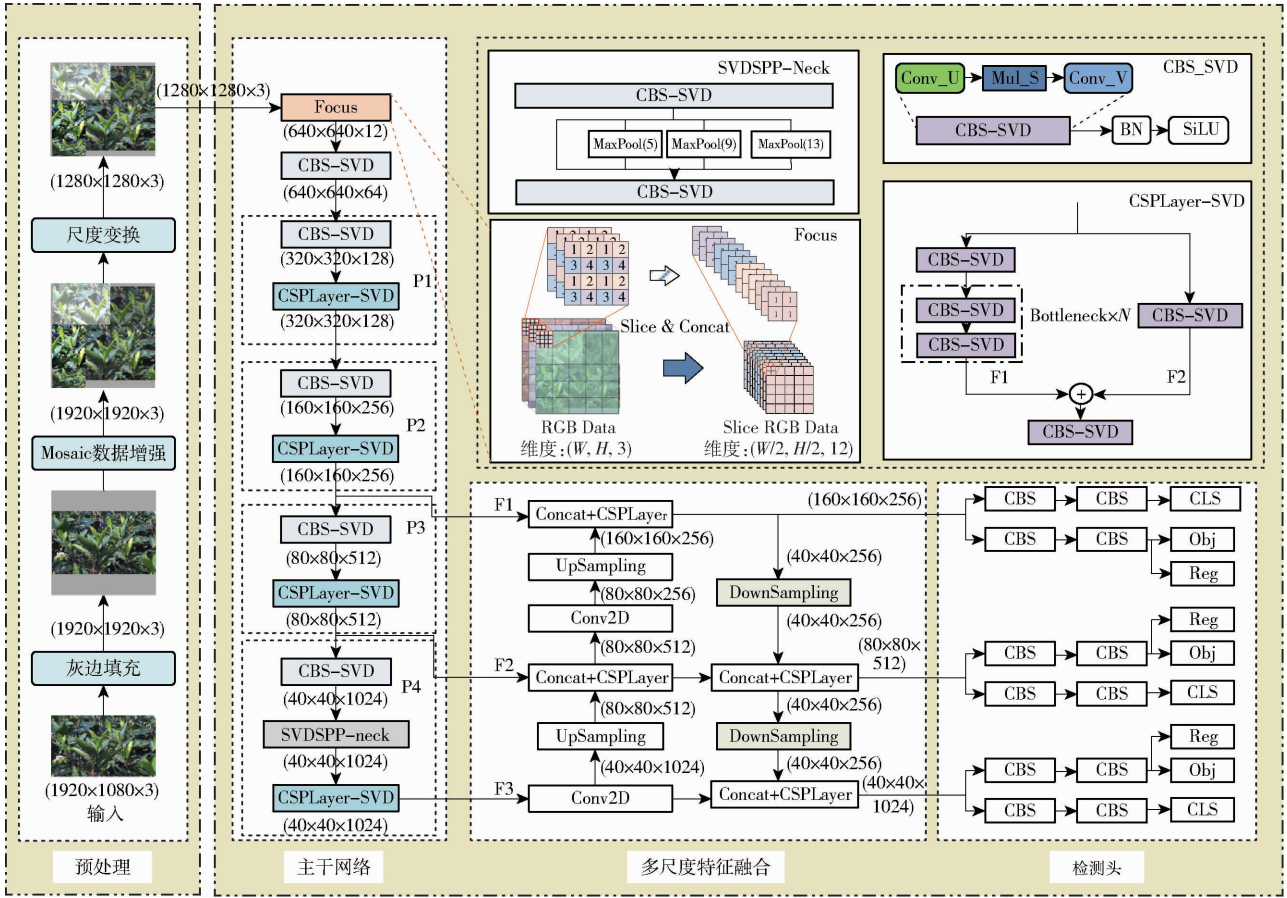


图 3 YSVD-Tea 目标检测算法

Fig.3 YSVD-Tea object detection algorithm

CSPLayer-SVD 结构作为一种残差结构, 由基本卷积块 Conv-SVD 与 N 个 Bottleneck 块组成, 其中 N 在不同的特征提取阶段取不同的数值。主干结构中从 $P1$ 到 $P3$ 的特征提取部分是通过交替连接 CBS-SVD 块和 CSPLayer-SVD 残差结构实现的。与之前的特征提取结构不同, $P4$ 特征提取块改用了 SVDSPP-Neck 模块, 这是一种由 CBS-SVD 卷积和池化结构组成的多尺度特征增强结构, 通过结合 CBS-SVD 结构和多尺度的池化操作, 该模块可以更好地将多尺度信息体现在提取

的特征中。与主干网络连接的是多尺度特征融合模块, 多尺度特征融合结构通过多组上下采样、特征连接和特征提取卷积块 CSPLayer 操作, 实现了不同尺度特征之间的融合。检测头采用解耦检测头的形式, 分别实现了分类和检测框回归任务。这种解耦结构使得不同任务的卷积权重能够分开学习, 从而在一定程度上降低学习难度, 提高网络学习效果。

1.2.3 使用遗传算法进行超参数调优

一组合适的训练超参数可以提高网络的总体性

能。为了寻找这样一组参数,通常通过人工搜索的方式设定多组实验,选出综合性能最好的一组作为最终的超参数。然而,人工超参数调优具有随机性,调参过程繁琐,且难以找到最优的超参数组合^[25]。因此,本文采用遗传算法进行超参数优化。通过对这些指标加权求和的方式构建出新的评价指标,用于评价目标检测网络的综合性能表现。在茶叶检测任务中,通常关注平均精度均值 (Mean average precision, mAP) 和 F1 值性能指标,它们相对于准确率和召回率能够更全面地反映网络性能。因此,为了能够均衡地衡量模型性能,赋予 mAP 和 F1 值指标较高的加权比率,同时赋予准确率和召回率较低的加权比率,公式为

$$F(x) = W_{\text{mAP}} I_{\text{mAP}} + W_{\text{F1}} I_{\text{F1}} + W_R I_R + W_P I_P \quad (2)$$

W_{mAP} 、 W_{F1} 、 W_R 、 W_P 分别表示网络性能测试指标。其中 I_{mAP} 、 I_{F1} 、 I_R 、 I_P 为性能指标对应的加权重,式(2)中的评价指标 $F(x)$ 用来评估个体适应度。设置 mAP 和 F1 值权重为 0.5,相对于仅使用 mAP 的情况,网络性能有所提升。然后随机尝试了多组加权重,最终选择表现最优的加权重,即 I_{mAP} 、 I_{F1} 、 I_R 、 I_P 分别为 0.44、0.44、0.06、0.06。

遗传算法流程图如图 4 所示,它包括个体定义、父代选择、突变操作、超参数约束、训练模型和性能记录、迭代进化等步骤。具体而言,个体定义是将一组超参数配置作为个体并将超参数的具体值作为个体的基因,它们包括了权重以及权重范围等信息。突变操作则是通过算法生成新的超参数配置,生成新的个体,从而达到探索新的超参数配置的作用。超参数约束是为了确保生成的新个体超参数在合理范围内。每个新个体都会使用适应度函数计算个体的适应度。然后通过父代选择,从之前的结果中选择一个父代个体,采用性能加权的方式选择优势基因。从而完成了一轮完整的遗传算法迭代。在满足终止条件前,不断重复上述步骤,生成多个个体,进行多代遗传算法进化,从而完成网络超参数选择。

2 结果与讨论

2.1 超参数调优结果

本文采用遗传算法对 YSVD - Tea 目标检测算法的训练超参数进行调优。在进行持续的变异、父代选择及迭代进化等步骤后,遗传算法最终成功迭代出 YSVD - Tea 算法的最佳超参数组合。迭代实验进行了 300 次计算,为了更清晰地呈现结果,将最后 30 次迭代结果以散点图的形式展示在图 5 中,图中红色标记突出显示参数搜索的最佳取值。

表 2 详细列出了遗传算法实验所需的初值、突变率以及取值范围等设置,这些设置在实验中发挥了关键作用。超参数的取值范围和迭代初值分别决定了超参数的搜索空间和搜索起点的位置。本文中采用了经验值来定义这些范围和初始值。此外,突变率是另一个重要因素,它用于控制每轮迭代中超参数的变化程度,从而影响搜索的随机性和速度。较高的突变率会增加搜索的随机性,使搜索更容易不稳定,而较低的突变率可能导致搜索陷入局部最优解。因此,根据不同超参数的潜在重要程度,实验过程中不同参数的突变率被设定为在 0 ~ 1 范围内的不同数值。通过多次迭代最终完成超参数搜索实验,搜索的最佳值见表 2。

测试集上的实验结果如表 3 所示。为深入比较 YSVD - Tea 算法在超参数优化前后的性能变化,设计了优化前后的对比实验。两组超参数训练过程中的指标变化如图 6 所示。图 6a 显示,优化后的召回率出现明显提升;图 6b 中的准确率也在第 110 个训练周期后出现了小幅提升。此外,如图 6c、6d 所示,优化后相较于优化前 box 损失值和 obj 损失值更低,收敛速度更快,证明超参数调优对算法性能提升的有效性。在调优过程中,适应度函数 $F(x)$ 的权重并未进行细粒度的遍历实验,因此可能并非最优的设置。尽管如此,经过调优后的性能提升显著。

2.2 YSVD - Tea 性能分析

训练完成后,YSVD - Tea 算法在测试集上的性能指标如表 3 所示。为更清晰地观察训练前后的变化,本文对中间特征进行了更详细的分析。在子空间 S 内,奇异值的大小反映了子空间 U 和 V 中的语义线索的重要程度^[23]。 S 中的奇异值越大,表明子空间 U 和 V 中的语义线索越为重要。因为主干网络的最后一层特征图包含了网络中所有层次的特征信息,其权重的变化能够比较全面反映模型特征信息。所以实验选取 YSVD - Tea 主干网络中最后一层特征进行分析。为更好地了解训练对模型提取的语义特征的影响,对训练前后奇异值发生显著变化

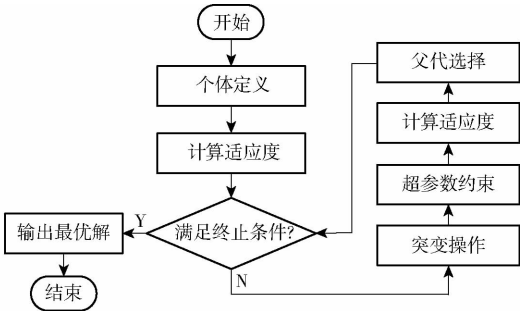


图 4 遗传算法流程图
Fig. 4 Genetic algorithm flowchart

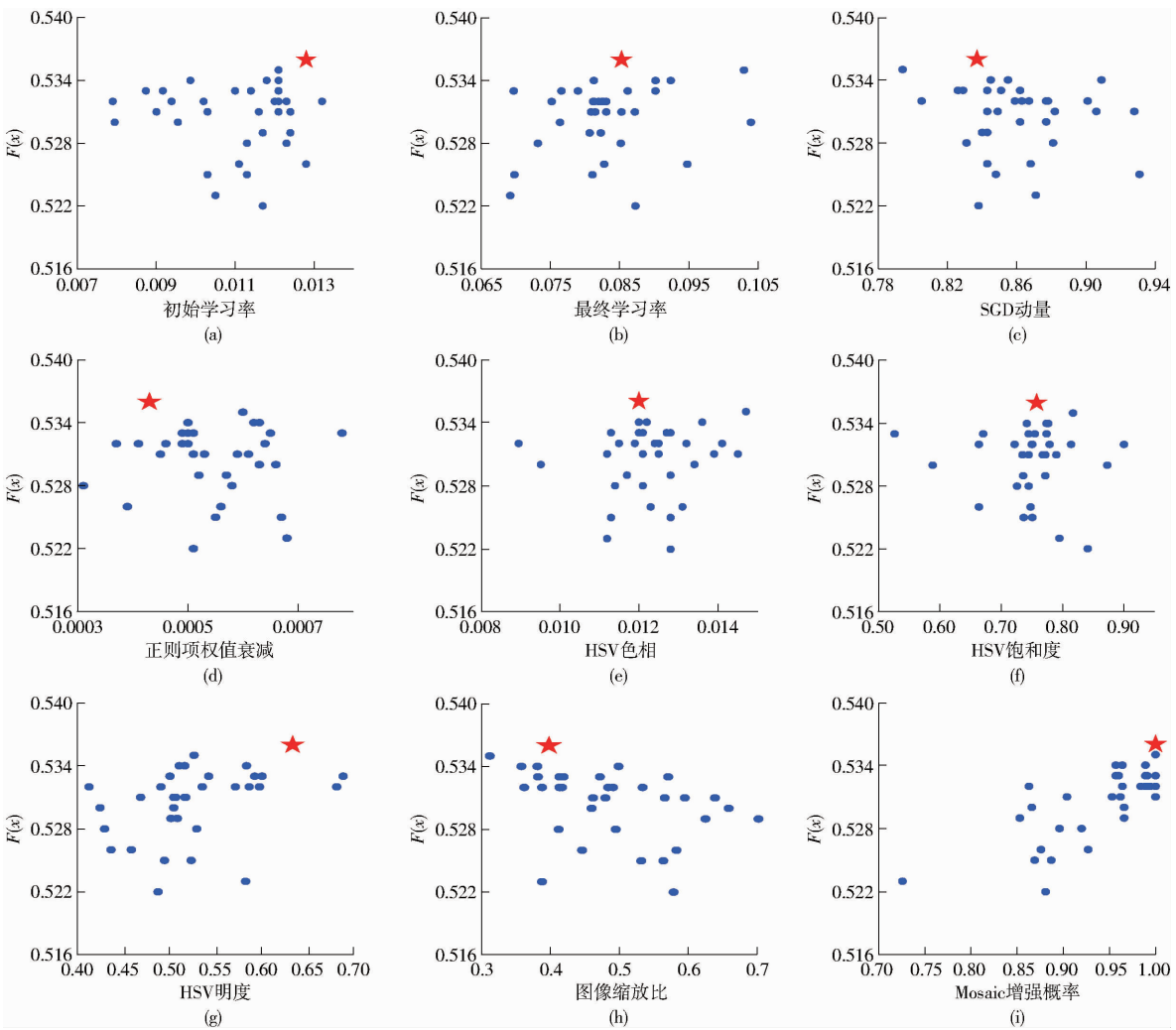


图5 超参数搜索散点图

Fig.5 Scatter plots of hyperparameter search

表 2 网络超参数

Tab.2 Hyper-parameters of model

参数名称	初值	取值范围	突变率	最佳值
初始学习率	0.01	0.000 01 ~ 0.1	1	0.012 8
最终学习率	0.1	0.1 ~ 1.0	1	0.085 3
SGD 动量	0.937	0.6 ~ 0.98	1	0.837
正则项权值衰减	0.000 5	0 ~ 0.001	1	0.000 43
图像缩放比	0.1	0 ~ 0.9	0.5	0.398
HSV 色相	0.015	0 ~ 0.1	0.8	0.012
HSV 饱和度	0.7	0 ~ 0.9	0.8	0.758
HSV 明度	0.4	0 ~ 0.9	0.8	0.634
Mosaic 增强概率	1.0	0 ~ 1.0	0.3	1.0

表 3 YSVD-Tea 算法超参数优化前后性能对比

Tab.3 Performance comparison of YSVD-Tea algorithm before and after hyperparameter optimization %

类别	mAP	召回率	精确率	F1 值
优化前	84.2	82.4	84.0	83.2
优化后	88.3	87.0	86.4	86.7

的语义线索进行了可视化分析。具体来说,训练过程中增大权重的部分往往与当前任务更相关,而减

小权重的部分则与当前任务的相关性降低。因此,本文统计了训练前后奇异值的变化,将权重变化划分为增加和减少两类,并将它们按照大小进行排序。通过可视化增幅最大的语义信息,以及减幅最大的语义信息(图 7)。由图 7 可以观察到,在训练过程中减少的语义线索倾向于背景区域,而增加的语义线索则主要集中在前景区域。这验证了在任务迁移过程中,奇异值所在的 S 子空间的变化与背景和前景特征权重的变化存在较强的相关性。总体而言,通过较少改变预训练模型语义信息的情况下完成任务训练,可以在一定程度上提高模型的泛化能力。这也使得 YSVD-Tea 模型能够在更好地保留预训练模型语义信息的同时,只通过在少量茶芽标注样本的训练就能实现较好的茶芽检测性能。

2.3 对比实验

2.3.1 不同规模数据集下 YSVD-Tea 与对比算法性能比较

为验证 YSVD-Tea 算法在少量标签数据下的优势,在完整数据集、1/2 数据集和 1/3 数据集上,

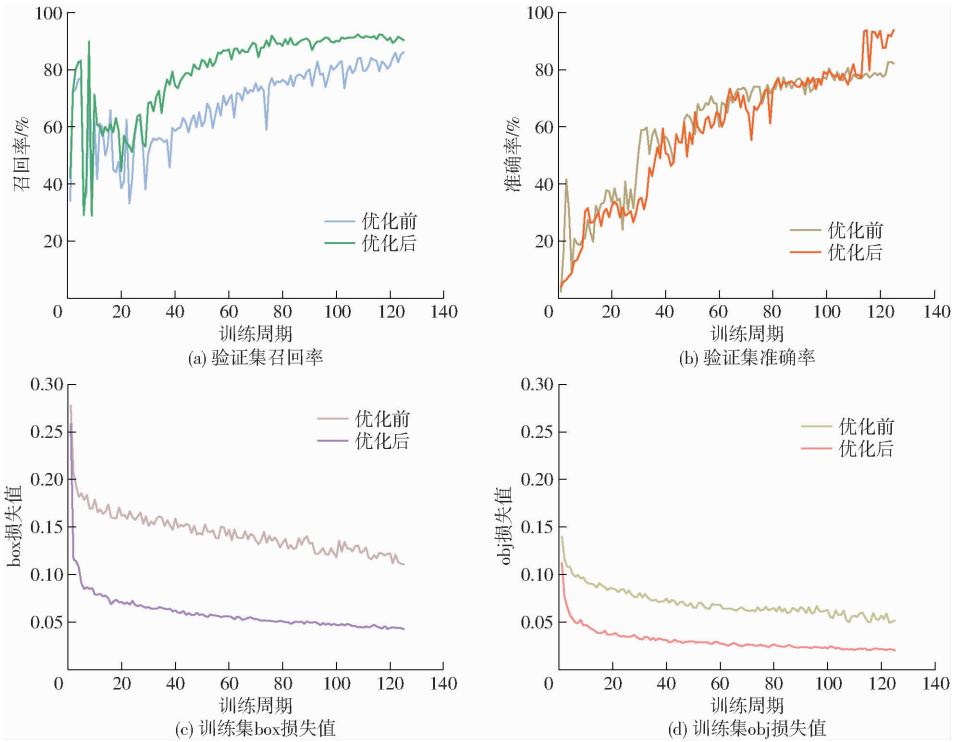


图 6 超参数优化对 YSVD - Tea 算法性能的影响

Fig. 6 Impact of hyperparameter optimization on performance of YSVD - Tea algorithm

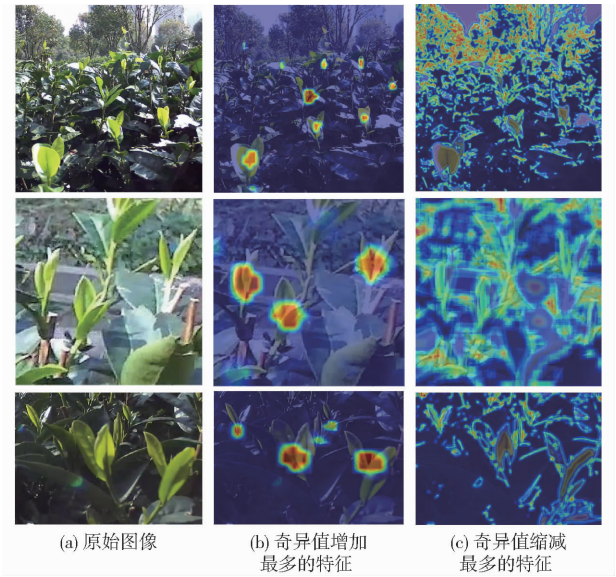


图 7 YSVD - Tea 语义信息可视化

Fig. 7 Semantic information visualization in YSVD - Tea

分别比较了本文算法与 YOLO v3^[26]、Faster R - CNN^[27]、YOLOX^[24] 算法的性能指标,结果如表 4 ~ 6 所示。

从表 4 可以观察到,在仅使用 1/3 的数据集情况下,YSVD - Tea 算法的 mAP 达到 60.9%。相较之下,YOLOX 的 mAP 仅为 40.6%,而 YOLO v3 和 Faster R - CNN 模型的 mAP 分别仅为 25.1% 和 21.9%。对比算法在这一数据规模下表现较差,而 YSVD - Tea 算法在有限样本上表现出显著的优势。当数据量扩大至 1/2 的数据集规模时,如表 5 所示,

表 4 模型在 1/3 数据集上的检测性能评估对比

Tab. 4 Detection performance comparison of models on 1/3 dataset

on 1/3 dataset				%
模型	mAP	召回率	精确率	F1 值
YOLO v3	25.1	19.0	38.3	25.4
Faster R - CNN	21.9	15.6	33.2	21.2
YOLOX	40.6	48.7	43.8	46.1
YSVD - Tea	60.9	61.3	57.7	59.4

表 5 模型在 1/2 数据集上的检测性能评估对比

Tab. 5 Detection performance comparison of models on 1/2 dataset

on 1/2 dataset				%
模型	mAP	召回率	精确率	F1 值
YOLO v3	40.6	48.7	43.8	46.1
Faster R - CNN	37.6	46.1	39.4	42.5
YOLOX	61.7	64.4	56.2	60.0
YSVD - Tea	76.2	74.3	69.4	71.8

表 6 模型在完整数据集上的检测性能评估对比

Tab. 6 Detection performance comparison of models on full dataset

on full dataset				%
模型	mAP	召回率	精确率	F1 值
YOLO v3	50.2	46.8	58.3	51.9
Faster R - CNN	62.5	52.4	70.6	60.2
YOLOX	74.1	61.2	81.2	69.8
YSVD - Tea	88.3	87.0	86.4	86.7

YSVD - Tea 算法的 mAP 显著提升至 76.2%。相较之下,YOLOX 的 mAP 仅为 61.7%,而 YOLO v3 和 Faster R - CNN 模型的 mAP 分别仅提

升至 40.6% 和 37.6%。如表 6 所示,当数据量扩展至完整数据集时,YSVD-Tea 算法的 mAP 进一步提升至 88.3%。与此同时,YOLOX 和 Faster R-CNN 算法的 mAP 也分别提升至 74.1% 和 62.5%。

图 8 以三维柱状图的形式更加清晰地展示了不同算法的性能对比。从图中可以看出,在 3 个不同数量的茶芽数据集实验中,YSVD-Tea 算法的性能一直是最佳的。特别是在有限样本情况下,与对比方法相比,YSVD-Tea 算法性能优势最大。随着数据集标签数量的增加,YSVD-Tea 算法的性能持续提升,同时性能指标依然保持良好的检测效果。这充分证明该算法在小规模数据下相对于对比算法的优越性,并且在数据集增加的情况下仍能够实现最佳的检测性能。

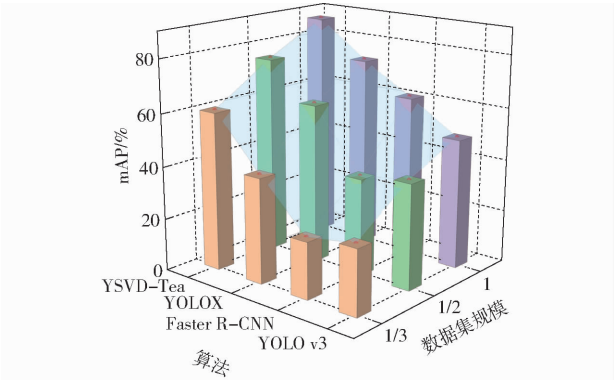


图 8 对比算法在不同规模数据集上的 mAP 性能
Fig.8 Comparative algorithms mAP performance on different dataset sizes

算法在训练集和测试集上的性能差距在一定程度上可以反映算法的泛化能力。为有效评估 YSVD-Tea 的性能提升原因,本文对比了相关算法在测试集和训练集上的性能,如图 9 所示。以 mAP 指标为例,在全数据集上,YOLO v3 在训练集和测试集上的性能差距为 36.4 个百分点,Faster R-CNN 为 17.8 个百分点,YOLOX 为 10.1 个百分点,而 YSVD-Tea 仅为 4.5 个百分点。在 1/2 数据集上,YOLO v3 在训练集和测试集上的性能差

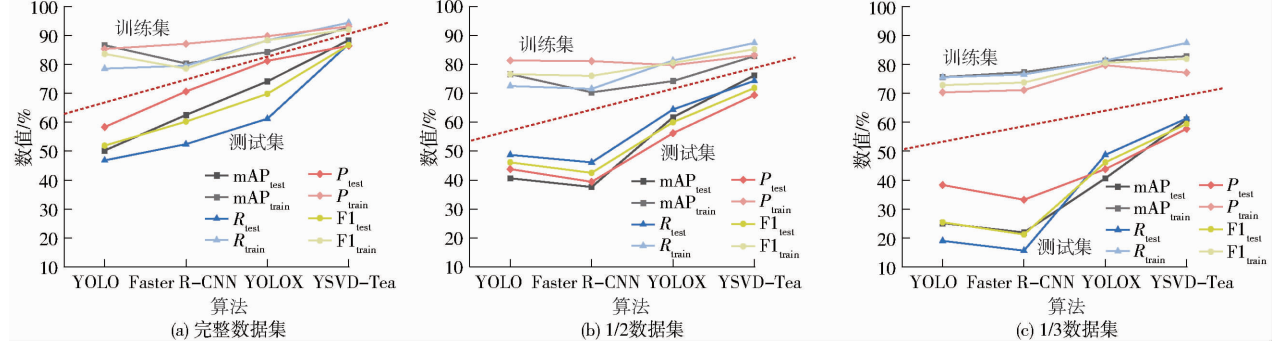


图 9 不同模型在训练集和测试集上的检测性能

Fig.9 Detection performance evaluation of model on training and test sets

距为 36 个百分点,Faster R-CNN 为 32.7 个百分点,YOLOX 为 12.5 个百分点,YSVD-Tea 仅为 6.6 个百分点。在 1/3 数据集上,YOLO v3 在训练集和测试集上的性能差距为 50.5 个百分点,Faster R-CNN 为 55.4 个百分点,YOLOX 为 40.6 个百分点,而 YSVD-Tea 仅为 21.9 个百分点。对比这 3 种规模的数据集,可以观察到随着数据减少,算法在训练和测试集上的性能差异变得更为明显。然而,YSVD-Tea 相对于对比算法在测试集与训练集上的性能指标更为接近,这表明 YSVD-Tea 具有更强的泛化能力。YSVD-Tea 采用了网络奇异值分解方法,在很大程度上保留了预训练模型学到的通用特征。在新的迁移任务中,该算法在学习茶芽新特征的同时,尽可能地保留了预训练模型的先验知识,以维持其良好的泛化能力。相比之下,其他对比方法由于训练样本数量较少,局限于训练集学到的特征,导致过度拟合于训练集信息,从而在测试集上表现出较差的性能,呈现出相对较弱的泛化能力。YSVD-Tea 算法在少量样本下相对于对比算法展现出更好的泛化能力,从而证明了其更适用于少量数据的训练场景。

2.3.2 密集和稀疏目标场景下 YSVD-Tea 与对比算法的检测性能比较

检测任务一般对目标遮挡较为敏感,因此选择了稀疏目标和密集目标两种场景进行更细致的观察,本次实验人工选取 63 个稀疏目标和 112 个密集目标进行对比实验,对不同算法的检测结果进行统计,如图 10 所示。

在稀疏目标检测方面,YSVD-Tea 模型展现出卓越的性能。设定置信度阈值为 0.5 时,该算法成功检测了 61 个目标,明显优于其他模型。具体而言,YOLO v3 成功检测了 31 个目标,Faster R-CNN 成功检测了 51 个目标,而 YOLOX 成功检测了 53 个目标。密集目标检测方面,YSVD-Tea 同样表现出色,成功检测了 101 个目标,领先于其他模型。另外,YSVD-Tea 目标置信度大于 0.7 的目标数也是

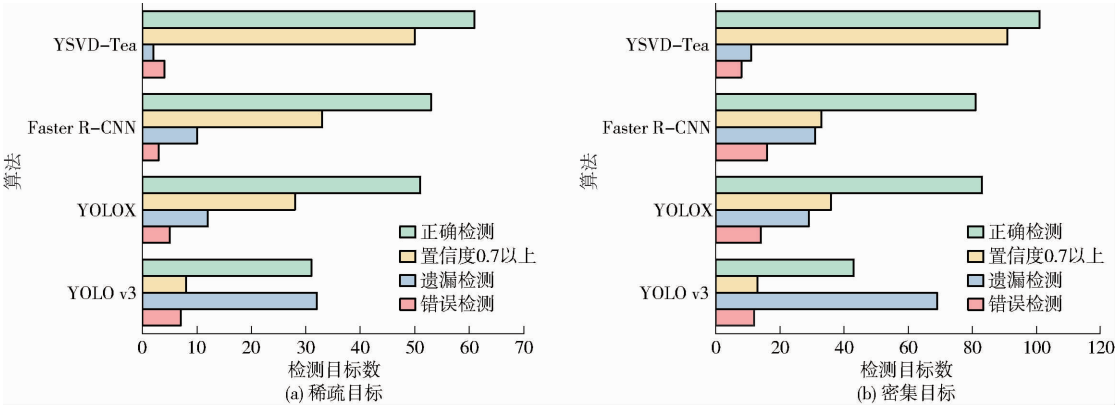


图 10 不同算法对密集和稀疏目标的检测结果

Fig. 10 Detection results of different algorithms for dense and sparse targets

最多的,共计 91 个,这也是所有模型中表现最佳的。综上所述,YSVD - Tea 算法在稀疏目标和密集目标两种场景下均表现出色。该算法不仅提高了检测准确性,还提供了高置信度的检测结果。检测结果如图 11 所示。在稀疏目标和密集目标两种场景下,

YSVD - Tea 算法呈现出相对准确的检测框,尤其是在密集目标情况下,YSVD - Tea 算法较少出现漏检的情形,而对比算法则在不同程度上存在漏检现象。这一结果间接验证了 YSVD - Tea 算法在这两个场景下的卓越性能。

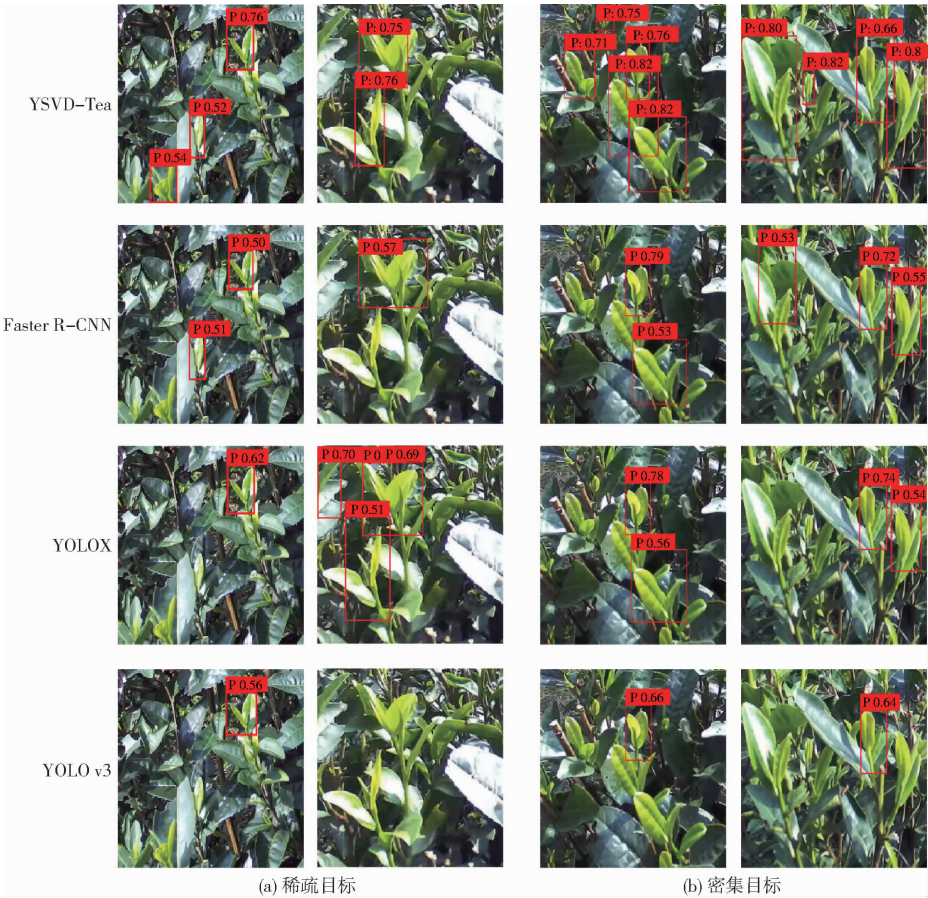


图 11 不同算法的检测结果

Fig. 11 Comparison of different algorithms in object detection

YSVD - Tea 算法表现出性能优势的可能原因在于,其对预训练模型中的语义信息的保留程度相较于其他对比算法更为显著。预训练模型具备出色的语义特征,其中包括对遮挡目标的识别经验。因此,YSVD - Tea 算法通过更有效地保留这些关键信

息,在密集目标和稀疏目标的场景下,相较于对比算法表现出更为卓越的检测性能。

2.3.3 基于少量样本的目标检测算法性能对比
为了深入评估有限样本情境下的目标检测性能,本文对比 3 种方法,分别是 ASPP - YOLO v5^[28]、

YOLO v5s - Ghost^[29] 和 E - YOLO^[30]。这些方法被用于解决田间苋菜识别、农作物虫害识别和水体遥感检测问题,同时这些方法也都在应对少量样本问题上做出了优化。为验证这些算法在有限样本情况下的性能,实验选用 1/3 规模的数据集。对比结果如表 7 所示,可以发现 YSVD - Tea 算法在各项性能指标上均达到最高,验证了其在有限样本目标检测任务中的有效性。由于 SVD 方法对预训练模型特征的较大程度保留,YSVD - Tea 网络在学习新特征的同时也保留了更多的预训练模型特征,从而保持了更好的泛化能力。因此,在有限样本目标检测任

表 7 基于少量样本的目标检测算法性能对比
Tab.7 Performance comparison of few-shot
object detection algorithms %

模型	mAP	召回率	精确率	F1 分数
ASPP - YOLO v5	43. 7	47. 5	48. 4	47. 9
YOLO v5s - Ghost	42. 8	52. 0	46. 2	48. 9
E - YOLO	47. 3	55. 2	52. 1	53. 6
YSVD - Tea	60. 9	61. 3	57. 7	59. 4

务中,YSVD - Tea 算法相较于其他方法表现较好,展现了其在有限样本情境下的目标检测适应能力。

3 结束语

本文提出的 YSVD - Tea 目标检测算法,基于奇异值分解原理,在小规模茶芽数据集上表现出色,mAP 为 88.3%,准确率为 86.4%,召回率为 87.0%,F1 值为 86.7%。该算法通过奇异值分解技术划分模型权重,有效平衡了适应茶芽检测任务和保留预训练模型语义信息的需求,提高了泛化能力。在少量标注样本下,YSVD - Tea 算法相较传统算法表现更为稳定,更适用于少标注样本的茶芽目标检测任务。此外,全局搜索算法对超参数的优化进一步提升了模型性能。经过实验验证,算法在密集和稀疏目标场景下都展现出色的识别与定位性能,具有较强的鲁棒性。研究成果为减少茶芽数据标注工作量、提高模型泛化能力等问题提供了一种可行的解决方案。

参 考 文 献

[1] 刘莫尘,褚镇源,崔明诗,等. 基于改进 YOLO v8 - Pose 的红熟期草莓识别和果柄检测[J]. 农业机械学报,2023,54(增刊 2):244 - 251.
LIU Mochen, CHU Zhenyuan, CUI Mingshi, et al. Red ripe strawberry recognition and stem detection based on improved YOLO v8 - Pose[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(Supp. 2):244 - 251. (in Chinese)

[2] 司永胜,孔德浩,王克俭,等. 基于 CRV - YOLO 的苹果中心花和边花识别方法[J]. 农业机械学报,2024,55(2):278 - 286.
SI Yongsheng, KONG Dehao, WANG Kejian, et al. Recognition of apple king flower and side flower based on CRV - YOLO [J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(2): 278 - 286. (in Chinese)

[3] 苗荣慧,李志伟,武锦龙. 基于改进 YOLO v7 的轻量化樱桃番茄成熟度检测方法[J]. 农业机械学报, 2023,54(10): 225 - 233.
MIAO Ronghui, LI Zhiwei, WU Jinlong. Lightweight maturity detection of cherry tomato based on improved YOLO v7[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(10): 225 - 233. (in Chinese)

[4] 宋彦,杨帅,郑子秋,等. 基于多头自注意力机制的茶叶采摘点语义分割算法[J]. 农业机械学报,2023,54(9):297 - 305.
SONG Yan, YANG Shuai, ZHENG Ziqiu, et al. Semantic segmentation algorithm based multi-headed self-attention for tea picking points[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(9):297 - 305. (in Chinese)

[5] 黄家才,唐安,陈光明,等. 基于 Compact-YOLO v4 的茶叶嫩芽移动端识别方法[J]. 农业机械学报,2023,54(3):282 - 290.
HUANG Jiakai, TANG An, CHEN Guangming, et al. Mobile recognition solution of tea buds based on Compact-YOLO v4 algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023,54(3):282 - 290. (in Chinese)

[6] XU W, ZHAO L, LI J, et al. Detection and classification of tea buds based on deep learning[J]. Computers and Electronics in Agriculture, 2022, 192: 106547.

[7] YAN L, WU K, LIN J, et al. Identification and picking point positioning of tender tea shoots based on MR3P-TS model[J]. Frontiers in Plant Science, 2022, 13: 962391.

[8] GUI Z, CHEN J, LI Y, et al. A lightweight tea bud detection model based on YOLO v5[J]. Computers and Electronics in Agriculture, 2023, 205: 107636.

[9] LI Y, MA R, ZHANG R, et al. A tea buds counting method based on YOLO v5 and Kalman filter tracking algorithm[J]. Plant Phenomics, 2023, 5: 0030.

[10] CHEN C, LU J, ZHOU M, et al. A YOLO v3-based computer vision system for identification of tea buds and the picking point [J]. Computers and Electronics in Agriculture, 2022, 198: 107116.

[11] 王梦妮,顾寄南,王化佳,等. 基于改进 YOLO v5s 模型的茶叶嫩芽识别方法[J]. 农业工程学报, 2023, 39(12): 150 - 157.
WANG Mengni, GU Ji'nan, WANG Huajia, et al. Method for identifying tea buds based on improved YOLO v5s model[J]. Transactions of the CSAE, 2023, 39(12): 150 - 157. (in Chinese)

- [12] ZHANG S, YANG H, YANG C, et al. Edge device detection of tea leaves with one bud and two leaves based on shuffleNetv2-YOLO v5-lite-E[J]. *Agronomy*, 2023, 13(2): 577.
- [13] CHEN Y T, CHEN S F. Localizing plucking points of tea leaves using deep convolutional neural networks[J]. *Computers and Electronics in Agriculture*, 2020, 171: 105298.
- [14] 崔金荣, 魏文钊, 赵敏. 基于改进 MobileNetV3 的水稻病害识别模型[J]. *农业机械学报*, 2023, 54(11): 217-224, 276. CUI Jinrong, WEI Wenzhao, ZHAO Min, et al. Rice disease identification model based on improved MobileNetV3[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(11): 217-224, 276. (in Chinese)
- [15] SUN Q, CHAI X, ZENG Z, et al. Multi-level feature fusion for fruit bearing branch keypoint detection[J]. *Computers and Electronics in Agriculture*, 2021, 191: 106479.
- [16] 汤文权, 陈灼廷, 王东桥, 等. 基于改进 YOLO v5 的皮蛋裂纹在线检测方法[J]. *农业机械学报*, 2024, 55(2): 384-392. TANG Wenquan, CHEN Zhuoting, WANG Dongqiao, et al. Crack detection method for preserved eggs based on improved YOLO v5 for online inspection[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2024, 55(2): 384-392. (in Chinese)
- [17] 刘世晶, 刘阳春, 钱程, 等. 基于 CycleGAN 和注意力增强迁移学习的小样本鱼类识别[J]. *农业机械学报*, 2023, 54(增刊 1): 296-302. LIU Shijing, LIU Yangchun, QIAN Cheng, et al. Recognition of small sample cultured fish based on CycleGAN and attention enhanced transfer learning[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(Supp. 1): 296-302. (in Chinese)
- [18] 李子茂, 徐杰, 郑禄, 等. 基于改进 DenseNet 的茶叶病害小样本识别方法[J]. *农业工程学报*, 2022, 38(10): 182-190. LI Zimao, XU Jie, ZHENG Lu, et al. Small sample recognition method of tea disease based on improved DenseNet[J]. *Transactions of the CSAE*, 2022, 38(10): 182-190. (in Chinese)
- [19] 袁培森, 宋进, 徐焕良. 基于残差网络和小样本学习的鱼图像识别[J]. *农业机械学报*, 2022, 53(2): 282-290. YUAN Peisen, SONG Jin, XU Huanliang. Fish image recognition based on residual network and few-shot learning[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(2): 282-290. (in Chinese)
- [20] LI S, SONG W, LI S, et al. Meta-retinanet for few-shot object detection[C]//British Machine Vision Conference (BMVC), 2020.
- [21] ZHANG G, CUI K, WU R, et al. PNPDet: efficient few-shot detection without forgetting via plug-and-play sub-networks[C]//2021 IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa: IEEE, 2021: 3822-3831.
- [22] WU A, ZHAO S, DENG C, et al. Generalized and discriminative few-shot object detection via SVD-dictionary enhancement [C]//Neural Information Processing Systems (NeurIPS), 2021.
- [23] SUN Y, CHEN Q, HE X, et al. Singular value fine-tuning: few-shot segmentation requires few-parameters fine-tuning[C]//Advances in Neural Information Processing Systems. Curran Associates, 2022: 37484-37496.
- [24] GE Z, LIU S, WANG F, et al. YOLOX: exceeding YOLO series in 2021[J]. *arXiv preprint arXiv:2107.08430*, 2021.
- [25] 余维, 李阳, 钟李红, 等. 基于改进实数编码遗传算法的神经网络超参数优化[J]. *计算机应用*, 2024, 44(3): 671-676. SHE Wei, LI Yang, ZHONG Lihong, et al. Hyperparameter optimization for neural network based on improved real coding genetic algorithm[J]. *Journal of Computer Applications*, 2024, 44(3): 671-676. (in Chinese)
- [26] REDMON J, FARHADI A. YOLO v3: an incremental improvement[J]. *arXiv preprint arXiv:1804.02767*, 2018.
- [27] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *Advances in Neural Information Processing Systems*, 2015, 28: 91-99.
- [28] 张继成, 侯郁硕, 郑萍, 等. 低数据集下基于 ASPP-YOLO v5 的苋菜识别方法研究[J]. *农业机械学报*, 2023, 54(增刊 2): 223-228. ZHANG Jicheng, HOU Yushuo, ZHENG Ping, et al. Method for amaranth identification based on ASPP-YOLO v5 model in low data set[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(Supp. 2): 223-228. (in Chinese)
- [29] 郭小燕, 于帅卿. 一种轻量级 YOLO v5s 农作物虫害目标检测模型[J/OL]. *南京农业大学学报*, <http://kns.cnki.net/kcms/detail/32.1148.S.20231204.1720.008.html>
- [30] 许泽宇, 沈占锋, 李杨, 等. 结合 E-YOLO 和水体指数约束的大篇幅影像水利设施检测[J]. *遥感学报*, 2022, 26(10): 2083-2093. XU Zeyu, SHEN Zhanfeng, LI Yang, et al. Detection of water conservancy facilities in large-format image combining E-YOLO algorithm and NDWI constraint[J]. *National Remote Sensing Bulletin*, 2022, 26(10): 2083-2093. (in Chinese)