DOI: 10.12398/j.issn.2096-7217.2024.02.005

基于改进YOLOv7的茶叶嫩芽识别模型研究

魏堂伟，张津诚，王晶，周庆燕*

（安徽农业大学信息与人工智能学院，安徽合肥，230036）

摘要：为了在复杂环境中有效识别茶叶嫩芽并在最大限度避免伤害茶树的情况下，提高智能化采摘精度，本研究针对传统目标检测算法在茶园中存在的检测精度低、鲁棒性差等问题，提出一种基于改进YOLOv7的茶叶嫩芽识别与检测的模型YOLOv7-tea，从而实现对茶叶嫩芽的快速识别与检测。首先，采集茶叶嫩芽图像并完成嫩芽图像的标注和数据增强，构建茶叶嫩芽数据集。其次，通过在YOLOv7主干网络的3个特征提取层中引入CBAM注意力机制模块，增强模型对特征提取的能力；采用SPD-Conv模块替换颈部网络下采样模块中的SConv模块，以减少小目标特征的丢失；通过EIoU损失函数优化框回归损失，提升预测框的准确率。最后，以茶叶嫩芽图像数据集为样本将其他目标检测模型与YOLOv7-tea模型进行对比试验，并对不同距离、不同角度拍摄的茶叶嫩芽图像进行识别效果检测。试验结果表明，YOLOv7-tea网络模型比YOLOv7模型的精确率（Precision，P）、召回率（Recall，R）和识别平均精度（mAP）值分别高出2.87、6.91和8.69个百分点，且模型的检测速度更快，在复杂背景下对茶叶嫩芽的识别与检测具有更高的置信度分数。该研究构建的YOLOv7-tea模型对小尺寸茶叶嫩芽的识别效果较好，减少了漏检和误检的情况，具有良好的鲁棒性和实时性，可为茶叶的产量预估和智能化采摘提供参考。

关键词：茶叶嫩芽；目标检测；CBAM注意力机制；自动化采摘；YOLOv7

中图分类号：S24；TS272.2 **文献标识码：**A **文章编号：**2096-7217（2024）02-0042-09

魏堂伟，张津诚，王晶，周庆燕. 基于改进YOLOv7的茶叶嫩芽识别模型研究[J]. 智能化农业装备学报(中英文)，2024，5(2)：42－50

WEI Tangwei，ZHANG Jincheng，WANG Jing，ZHOU Qingyan. Study of tea buds recognition and detection based on improved YOLOv7 model [J]. Journal of Intelligent Agricultural Mechanization，2024，5(2)：42－50

0 引言

近年来，中国的茶园面积和茶叶产量均快速增长[1]。作为主要的茶叶贸易国[2]和消费大国[3]，随着茶叶需求的增长，茶叶产量成为了人们关注的对象，而衡量茶叶产量需要确定茶叶嫩芽的数量。精准识别茶叶嫩芽是实现嫩芽计数的基础[4-5]，随着基于深度学习的目标检测技术的发展，该领域已取得了一定成果[6]。

许多专家学者展开了对茶叶嫩芽识别与检测的相关研究。SHAO等[7]采用改进的K-means算法对茶芽进行识别，获得了良好的效果；YAN等[8]提出了一种基于轻量级卷积神经网络MC-DM的茶叶分割和采摘点定位方法；ZHANG等[9]提出了一种基于MDY7-3PTB的模型来实现对茶叶嫩芽的分割及采摘点定位，模型改进后具有很强的鲁棒性；CHENG等[10]通过训练改进的Mask R-CNN网络对图像中的茶芽进行检测，可以检测到茶叶嫩芽区域中的关键点；LI等[11]改进了YOLOv4算法，可以准确地检测出茶叶嫩芽。龙樟等[12]提出了一种基于超绿特征的图像分割方法，对算法优化后实现了对嫩芽的识别。杨福增等[13]利用图像处理技术提取茶叶的颜色和形状特征，实现对嫩芽的识别。王梦妮等[14]通过改进YOLOv5s网络模型，提高了茶叶嫩芽检测的精度。俞龙等[15]通过改进YOLOX模型准确地对茶叶嫩芽进行识别，降低漏检率和错检率。

以上的研究方法主要依赖于茶叶的颜色和形状特征，在复杂环境下容易受光照和枝干遮挡影响，导致准确率降低。一些方法仅适用于特定品种或条件，尽管已取得一定成果，但在提高泛化能力、检测速度和降低模型复杂度方面仍需改进。为了精确识别茶叶嫩芽并实现其自动化采摘，本研究构建了一种改进的YOLOv7的茶叶嫩芽识别模型YOLOv7-tea。该模型

收稿日期：2024-02-01 修回日期：2024-03-20

基金项目：安徽省科技计划项目（2022MKS12）

第一作者：魏堂伟，女，1997年生，安徽六安人，硕士研究生；研究方向为人工智能。E-mail：1784725066@qq.com

*通信作者：周庆燕，女，1988年生，安徽安庆人，博士，讲师；研究方向为人工智能。E-mail：swallowchou@126.com

着重于提升网络对小目标的特征提取能力，提升对茶叶嫩芽的识别精度。

1　材料与方法

1.1　试验材料

本研究使用的茶叶嫩芽数据集包括迎霜、中茶 108 和中白 1 号等品种的茶叶嫩芽图像，这些图像于 2023 年 4 月 8－9 日在中国江苏省南京市江宁区横溪街道创业园区内的高标准茶园内进行采集。图像采集设备为 iPhone XR 手机，所有图像均以 JPG 格式保存；拍摄时间为每天的 8:00－17:00。

进行图像采集时以嫩芽为前景，嫩芽以外的信息为背景，拍摄角度与竖直向上方向呈 0°～90°夹角，拍摄距离为 10～60 cm。共采集了 1 200 张原始图像，用于构建试验所需的数据集。这些图像包含不同背景复杂度、近远景以及单张图像中目标数量不同的情况，如图 1 所示。图 1(a)(b)中，两个不同角度拍摄的图像中存在目标数量较少的茶叶嫩芽，同时没有太多老叶干扰，背景复杂程度较低；图 1(c)(d)中，不同距离拍摄的图像中存在较多的嫩芽目标分布密集和嫩芽存在遮挡情况，增加了模型的检测难度，在进行嫩芽检测时可能出现漏检情况。



图 1　不同角度和距离拍摄的茶叶嫩芽

Figure 1　Pictures of tea buds taken from different angles and distances

1.2　数据集标注与构建

在本研究中使用 LabelImg 标注软件对采集的茶叶图像进行了人工标注，考虑到真实采摘标准，在进行人工标记时，只针对图像中的一芽一叶进行标记，确定其为茶叶嫩芽。通过绘制矩形框来标记茶叶嫩芽的位置，并将类别设置为“tea”。完成上述操作后，本研究剔除了前景或背景模糊、无嫩芽和无茶梗的茶叶图像，将经过筛选后的 1 060 张茶叶图像按照 8:1:1 的比例分为训练集、测试集和验证集，3 种数据集分配数目分别为 848、106、106 张。

由于基于目标检测算法构建的模型需要大量的训练样本，而图像的采集受时间限制，因此当训练样本较少时，容易出现过拟合现象，从而影响模型的训练效果。为了防止过拟合并提高模型的泛化能力，本试验采用旋转、翻转、颜色增强等方式对茶叶图像进行数据增强处理，以增加训练集的图像数量。具体的数据增强方式如图 2 所示。

经过数据增强后，茶叶嫩芽数据集的训练集包含 1 696 张图像，测试集和验证集各包含 106 张图像，最终茶叶嫩芽数据集总样本数量达到 1 908 张。在训练集中，“tea”标签共被标记了 4 065 次。本研究以整理后的茶叶嫩芽图像为研究对象，构建了茶叶嫩芽数据集，为后续的茶叶嫩芽识别及检测方法研究提供了理论支持和数据基础。



图 2　茶叶嫩芽数据增强

Figure 2　Tea bud date enhancement

2　改进 YOLOv7 的茶叶嫩芽识别方法

2.1　YOLOv7 算法

YOLO 系列算法作为单阶段目标检测算法，以其较小的网络结构和较快的检测速度而著称。2022 年提出的 YOLOv7[16] 算法相比之前的 YOLO 算法在精度与速度等方面都有所提升[17]，因此，本文选用 YOLOv7 作为基础模型进行研究。

YOLOv7 模型的构成包含 4 个部分：Input 输入层、Backbone 骨干层、Neck 颈部层和 Head 头部层。模型的 Input 层负责对输入的图像进行预处理，统一调整图像尺寸为 640 像素×640 像素，以满足 YOLOv7 模型的输入要求；预处理后的数据在 Backbone 层经过处理，以提取各种尺寸目标的特征信息；提取到的特征经过 Neck 层特征融合处理，生成 3 种不同尺寸（大、中、小）的特征图；最终，融合后的特征被送入 Head 部分，经过重参数化结构 REP 层和 Conv 层进行最终处理，以输出检测结果。

2.2　CBAM 注意力机制

近年来，注意力机制在深度学习的众多领域中得到了广泛应用，该机制通过对特定位置施加权重来聚焦于关键部分，以此筛选出重要信息并降低次要信息的干扰。考虑到茶叶嫩芽识别与检测需要同时关注其语义特征和空间位置，本研究在 YOLOv7 算法的 Backbone 部分的 3 个特征提取层中引入 CBAM (convolutional block attention module) 注意力机制模块，CBAM[18] 是在 SENet 基础上进一步发展的混合域注意力机制。它由通道注意力模块和空间注意力模块组成，这两个模块可以单独或串联使用，具备通用性高

与轻量化这两大特点。算法的过程主要分为两步:首先,通过通道注意力机制筛选出特征图中的关键特征;其次,这些特征被输入到空间注意力模块以进一步确定目标的焦点区域。具体来说,通道注意力机制负责识别特征图中的关键通道,而空间注意力机制则通过压缩特征向量的通道维度并利用卷积来计算空间维度的权重,随后使用Sigmoid函数对特征进行加权。这种空间与通道注意力机制的结合使得模型在训练过程中能够更专注于对提升精度有重要影响的通道和区域。CBAM的结构如图3所示。

通过YOLOv7网络的各个特征提取层中引入了CBAM注意力机制这项改进策略,既提升了模型的主要特征提取效率,又强调了茶叶嫩芽的重要特征,同时减少了不重要特征的影响[19]。虽然引入CBAM注意力机制增加了一定的计算复杂度,但由于其对特征的有效提取,整体上提升了模型的性能。该改进算法的核心原则是在保证模型能够正常检测目标的基础上,提高目标的定位能力和检测准确率。引入CBAM注意力机制(红色区域)后的YOLOv7模型网络结构如图4所示。

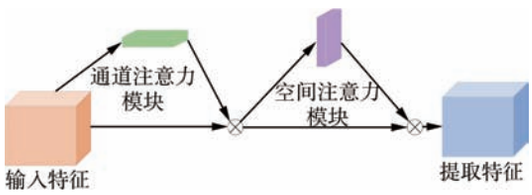


图3 CBAM注意力机制结构图

Figure 3 CBAM structure

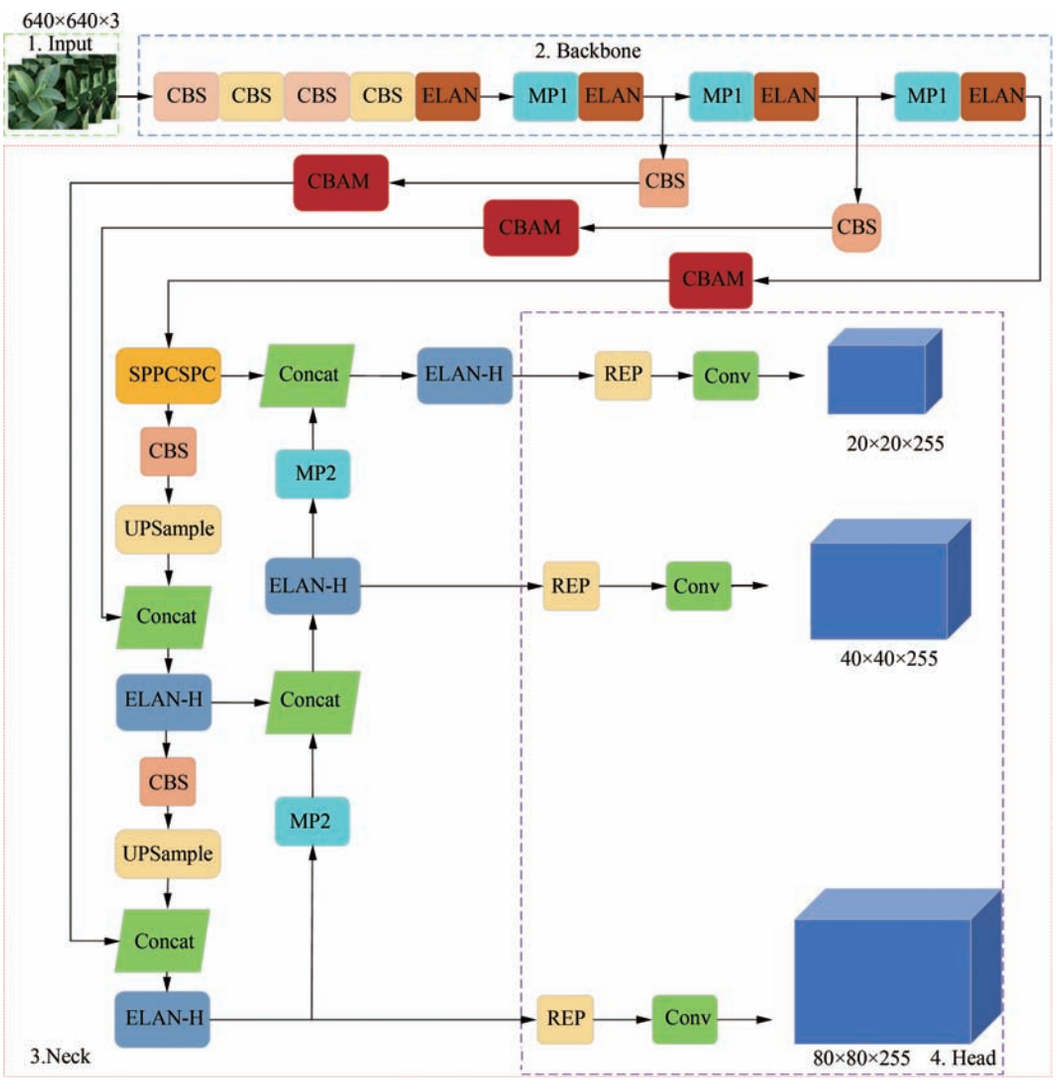


图4 引入CBAM注意力模块后的YOLOv7网络结构图

Figure 4 YOLOv7 network structure diagram after introducing CBAM attention module

2.3 引入SPD-Conv下采样模块

SPD-Conv[20]模块包括一个空间到深度(SPD)层和一个非跨步卷积层,SPD层用于下采样,而非跨步卷积层用于进一步提取特征,是一个专为低分辨率图像和小物体检测而设计的卷积神经网络构建模块。SPD-Conv模块可以在不失去可学习信息的情况下对

特征图进行下采样，摒弃了现在广泛使用的步长卷积和池化操作，其结构如图 5 所示。

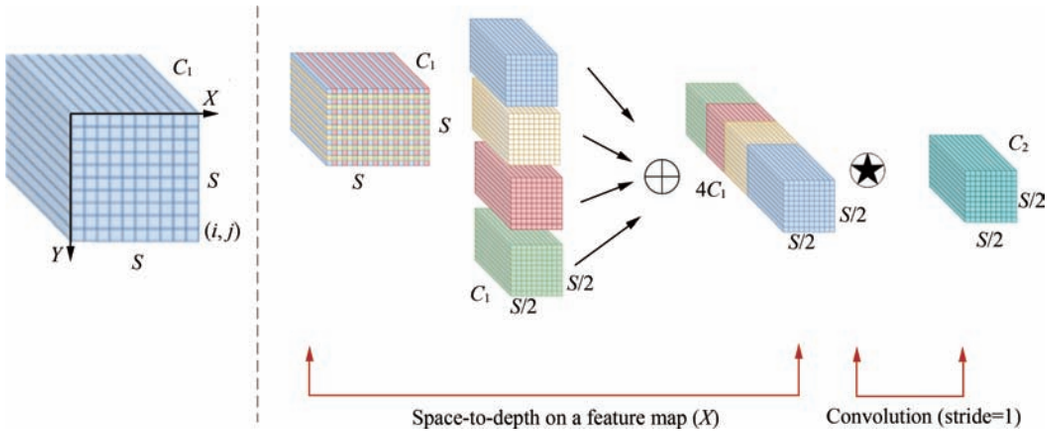


图 5　当 scale=2 时 SPD-Conv 的图解

Figure 5　Illustration of SPD-Conv when scale=2

从结构图可以看出，SPD-Conv 的结构较为简单。其中 SPD 层实现下采样，保留了通道维度中的所有信息，从而减少信息丢失，而非跨步卷积层用于进一步提取特征。与传统的 Depthwise Convolution 相比，SPD-Conv 通过其新颖的设计，能够更有效地提取小目标特征，提高检测性能。具体而言，SPD-Conv 通过引入空间注意力，可以更加注重目标的空间分布，有助于提取小目标的微小特征；同时，它在保持高性能时具有更少的参数数量，减少了模型的复杂性，并提高了整体模型的效率。

原 YOLOv7 网络结构中的 Neck 部分多次应用 SConv 来过滤冗余信息，但这可能会过滤掉茶叶嫩芽的关键特征。本研究为了降低微小目标特征的丢失，对 Neck 网络中 MP 结构进行了改进：将左分支中标准卷积 SConv 替换为 SPD-Conv，这一改进利用了 SPD 层的子图切分原理，将小目标全局空间信息有效地映射到通道维度，从而在光线昏暗和模糊场景中保留了更多小目标的特征信息。此外，在 SPD 层之后增加了卷积操作，以调整通道数，使其与网络的右分支保持一致。改进后的下采样结构见图 6。

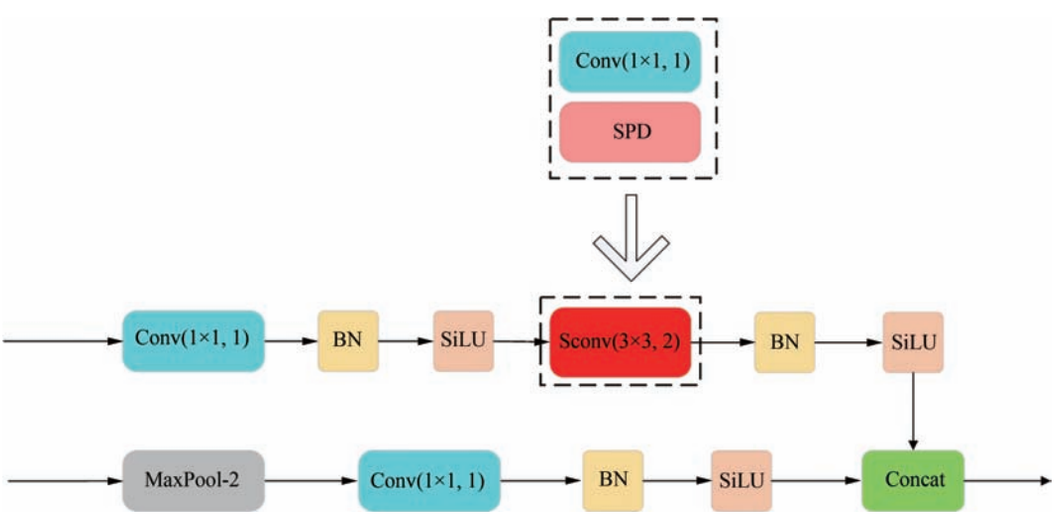


图 6　改进后 MP2 结构图

Figure 6　Improved MP2 structure diagram

2.4　替换损失函数

损失函数用于衡量模型预测结果与真实结果之间的差异。当对已标注的数据集进行多次迭代训练时，损失函数会逐步降低并最终达到稳定状态，此时损失函数的值最小，此刻可确定模型主干网络的参数。在训练过程中，计算模型的损失值是必要的，较低的损失值意味着模型预测值与真实值更为接近，从而反映出模型具有更高的稳定性和准确性。

在原 YOLOv7 中，定位损失采用 CIoU_Loss 函数计算，CIoU_Loss 损失函数能够对两个框未重叠的部分进行优化，从而加快了预测框的回归收敛速度。然而，当两个框的宽高比相似但不完全相等时，由于它们是相对值，可能会造成预测框在回归时宽度和高度无法同步增减，这会在一定程度上影响到模型的精

度。针对该问题,本研究中使用EIoU_Loss[21]函数替换了原YOLOv7中的损失函数。EIoU_Loss函数的主要贡献有两个,一是将宽高比的损失分开计算,二是解决了难易样本不均衡的问题。其计算公式如式(1)所示。

$$EIoU_Loss = 1 - CIoU = 1 - \left[IoU - \frac{P^2(b, b^{gt})}{C^2}\right] - \alpha v \quad (1)$$

式中:b,b^{gt}——预测框与真实框的中心点坐标;

P——两个中心点间的距离;

C^2——预测框与真实框的最小外接矩形的对角线距离;

αv——拟合预测框与实际框之间的纵横比。

EIoU认为简单的边界框回归任务不需要用过大的权重来学习,而复杂的边界框回归任务则需要更大的权重,以加快收敛速度并提高定位精度。此外,EIoU损失函数通过最小化真实框和预测框的宽度与高度之差,准确地衡量了重叠区域、中心点和边长的差异,有效地解决了样本难易程度不平衡的问题。为了进一步提高YOLOv7算法在茶叶嫩芽检测中的精度,本研究将YOLOv7网络中的CIoU损失函数替换为EIoU损失函数,并进行网络训练。

2.5 模型训练与算法性能评价

2.5.1 模型训练

本研究中的模型训练和测试使用Python3.7进行程序编写,并在配备4核CPU和NVIDIA A4000显卡的计算机上进行训练。试验环境的软件操作系统为Linux,深度学习框架为Pytorch1.7.1。

所有试验均在此一致的软硬件环境下完成。安装好相关环境配置后,使用构建的茶叶图像数据集进行了GPU训练,设置了周期(epoch)为300轮的训练,每批次输入16张图片(batch size),选择Adam优化器,初始学习率设为0.001,输入图片分辨率为640像素×640像素,所有模型的评估指标均在IoU阈值为0.5的条件下进行计算。

2.5.2 评价指标

为评估茶叶嫩芽识别与检测模型的性能,本研究采用精确率(Precision, P)、召回率(Recall, R)、平均精度均值(mAP)和$F1$值作为识别精度评价指标。精确率表示被正确预测为该类别的样本数占所有被预测为该类别的样本数的比例,其计算公式如式(2)所示。

$$P = \frac{TP}{TP + FP} \times 100\% \quad (2)$$

召回率是指模型正确识别为正类的样本数占所有正类样本数的比例。在茶叶嫩芽检测的场景中,它表示所有嫩芽中被正确检测出的比例,计算公式如式(3)所示。

$$R = \frac{TP}{TP + FN} \times 100\% \quad (3)$$

AP是综合衡量了在不同召回率情况下的精确率,而mAP平均精度均值则是综合考虑了所有检测类别的AP值均值,mAP的计算公式如式(4)所示。

$$mAP = \frac{\sum AP}{N} \times 100\% \quad (4)$$

$F1$得分是精确率和召回率的调和平均值,是衡量模型精确性和稳健性的指标。$F1$得分越高,说明模型在精确率和召回率之间达到了较好的平衡,其计算公式如式(5)所示。

$$F1 = \frac{2 \times P \times R}{P + R} \quad (5)$$

式中:TP——模型识别茶叶嫩芽的正确数量;

FP——模型识别茶叶嫩芽的错误数量;

FN——模型遗漏识别茶叶嫩芽的数量;

AP——茶叶嫩芽的平均精度;

N——模型能检测出的目标种类数量。

考虑到茶叶嫩芽检测的实际应用中对实时性的要求,本研究还将每秒检测帧数(FPS)作为模型识别速度的评价指标,通常,30帧/s的FPS被视为可以满足实时检测的需求。上述指标可以综合反映模型的实用性和效率。

3 试验结果与分析

3.1 消融试验

本研究设计了4组消融试验,以验证YOLOv7-tea模型的有效性以及各个改进策略对模型性能的影响。改进策略分别为:首先,在YOLOv7的3个特征提取层中引入CBAM注意力机制模块;其次,为了进一步优化模型,在颈部网络的下采样MP2结构的左分支中,将标准卷积层(SConv)替换为空间到深度卷积层(SPD-Conv);最后,使用EIoU损失函数替换原YOLOv7中的CIoU损失函数,来优化框回归。

本研究采用相同的数据集、试验环境和参数配置进行了试验,将mAP和识别精确率作为模型检测精度的评价指标,将FPS作为模型检测速度的评价指标,结合这3个指标全面地评估模型的性能表现,验证所用改进方法对模型性能的影响。试验结果汇总在表1中,其中“√”表示采用该模块对模型进行改进。

表 1　消融试验结果

Table 1　Results of ablation test

模型	CBAM	SPD-Conv	EIoU_Loss	*mAP*/%	*P*/%	*FPS*/(帧·s^{-1})
YOLOv7				71.48	87.45	56.6
YOLOv7-2	√			78.28	88.89	55.7
YOLOv7-3	√	√		79.95	90.27	55.4
YOLOv7-tea	√	√	√	80.17	90.32	57.2

由表 1 可知：原始的 YOLOv7 算法，其 *mAP* 为 71.48%。试验 YOLOv7-2 在原始 YOLOv7 的基础上增加了 CBAM 注意力机制模块，*mAP* 提升了 0.068。YOLOv7-3 试验在 YOLOv7-2 的基础上添加了 SPD 下采样模块，*mAP* 提升 0.016 7，这表明，引入注意力机制的改进下采样模块对模型性能有显著的提升。YOLOv7-tea 即本文方法在 YOLOv7 的基础上优化了 EIoU 损失函数，*mAP* 比 YOLOv7 提高了 0.086 9；该模型的精确率 Precision 也得到了提升，相较于 YOLOv7 增加了 0.028 7。检测速度方面，YOLOv7-tea 模型与原 YOLOv7 模型相差不大，由于采取 SPD-Conv 模块替换了颈部网络下采样模块中的 SConv 模块，增加了一定的计算量，但是 FPS 仍可以满足实时检测的需求。

综上所述，本文提出的 YOLOv7-tea 算法在检测精度方面相较于原始 YOLOv7 算法实现了显著提升，同时保持了较高的推理速度，证明了该算法在实时茶叶嫩芽识别任务中的有效性和实用性。

3.2　与其他算法对比分析试验

训练完成后，本研究对基于 YOLOv7-tea 算法训练得到的模型进行了可视化处理，从而获得了该模型在训练集和验证集上的损失值信息，如图 7 所示。

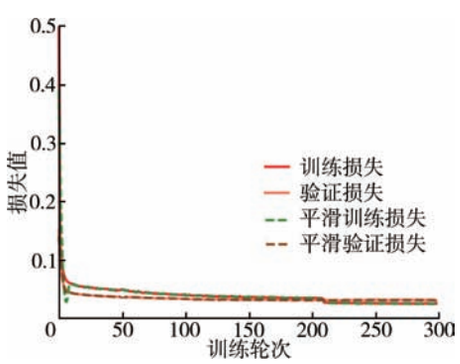


图 7　YOLOv7-tea loss 曲线

Figure 7　YOLOv7-tea loss curve

从图 7 可以看出，模型的损失值随着训练轮数的增加而呈现逐渐下降的趋势，并在大约 250 轮训练后达到稳定状态。本研究选择损失值最小的网络参数作为茶叶嫩芽的识别模型，最终模型的精确率达到了 90.32%，满足了茶叶嫩芽检测的需求。

为展示 YOLOv7-tea 算法的优越性，本研究进行了一组对比试验。将改进后的 YOLOv7-tea 算法与 Faster R-CNN（两阶段）、YOLOv5、YOLOX、YOLOv7（单阶段）等主流算法在相同的试验环境和数据集上进行了 300 轮训练，这些算法在茶叶嫩芽识别任务上的性能指标汇总在表 2 中。在综合比较准确率和检测时间后，可以看出 YOLOv7-tea 算法在效率和准确性方面展现出明显的优势。

表 2　不同检测算法的试验结果对比

Table 2　Comparison of experimental results from different detection algorithms

模型	*F*1 值	*mAP* /%	*R* /%	*FPS* /(帧·s^{-1})	*P* /%
Faster R-CNN	0.68	51.94	63.52	21.3	72.12
YOLOv5	0.59	56.48	48.77	49.0	73.46
YOLOX	0.65	59.61	60.25	55.3	71.36
YOLOv7	0.76	71.48	66.77	56.6	87.45
YOLOv7-tea	0.81	80.17	73.68	57.2	90.32

为验证 YOLOv7-tea 模型的实际识别性能，本研究在相同的试验条件下使用 106 张茶叶嫩芽图像对 Faster R-CNN、YOLOv5、YOLOX、YOLOv7 和 YOLOv7-tea 模型进行了测试，识别结果如图 8 所示，其中从左至右依次展示了不同模型对于 30°角、90°角、近距离和远距离拍摄的茶叶嫩芽图像的识别效果。

识别结果显示，Faster R-CNN、YOLOv5 和 YOLOX 模型在茶叶嫩芽的识别上效果较差，存在较多漏检和误检情况，错误地将非茶叶嫩芽的区域识别为茶叶嫩芽。YOLOv7 模型表现优于上述 3 种算法，但与 YOLOv7-tea 模型的置信度相比较低，识别效果一般且检测精度不够稳定。

因此，可以得出结论 YOLOv7-tea 模型在泛化性和鲁棒性方面均优于其他主流模型，在检测效果方面展现了更优异的检测性能，更适用于茶叶嫩芽的精准识别任务。

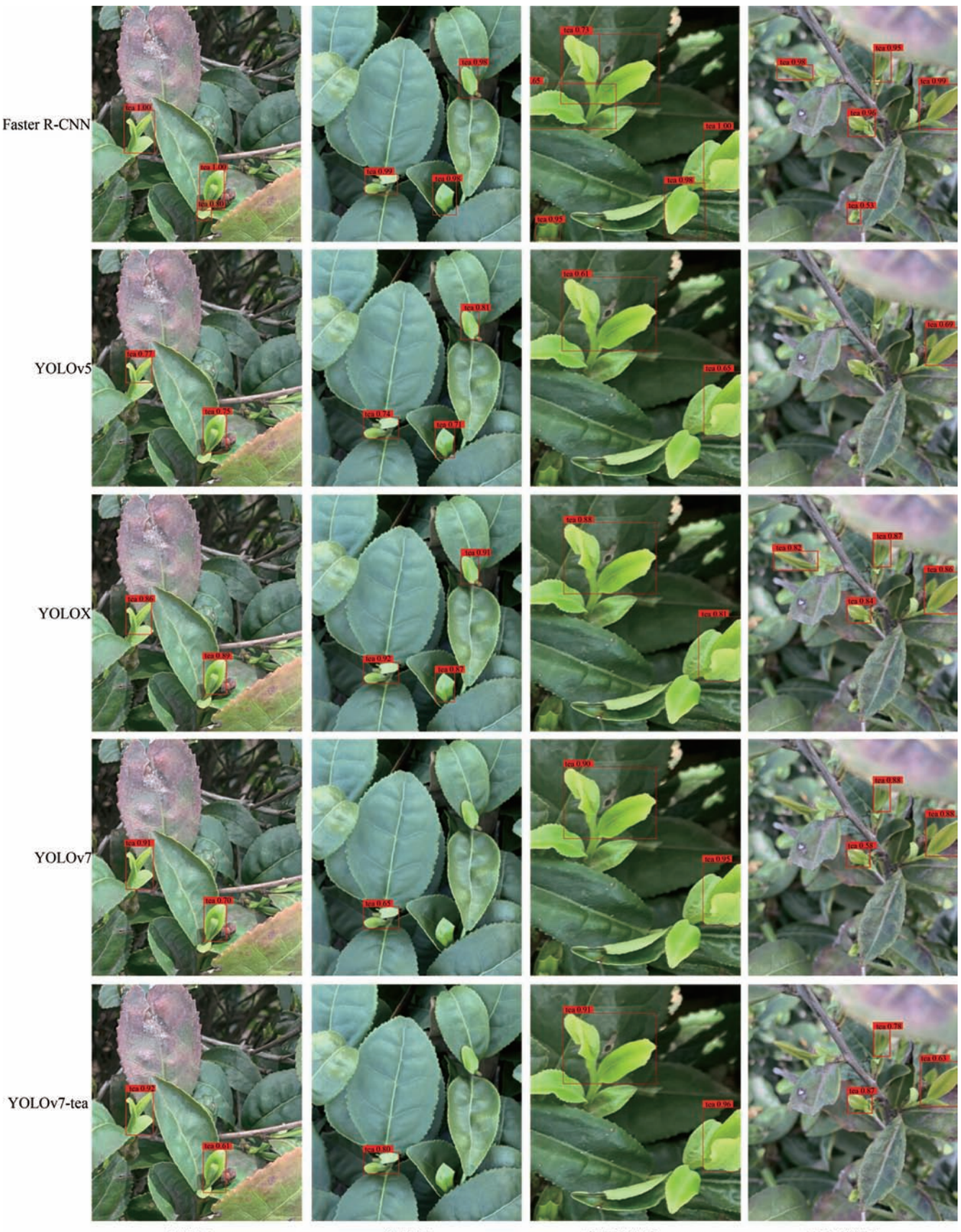


图 8　五种目标检测算法对茶叶嫩芽的检测效果

Figure 8　The detection effect of five target detection algorithms on tea buds

4　结论

本文针对茶园复杂环境下茶叶嫩芽识别精确率低、鲁棒性差的缺点，提出了一种基于改进YOLOv7的茶叶嫩芽识别与检测的模型YOLOv7-tea，可对不同角度和不同距离的茶叶嫩芽图像进行识别任务，为后续的智能采茶工作提供理论基础。具体结论如下。

1）在YOLOv7的主干网络特征提取层中引入了CBAM注意力机制模块；将颈部网络的MP2结构左分支中的标准卷积层（SConv）替换为空间到深度卷积层（SPD-Conv）；采用EIoU损失函数来优化边界框回归损失。改进后的YOLOv7-tea模型检测精度高，推理速度快，可以更准确地检测茶叶嫩芽目标。

2）经过综合评估YOLOv7-tea和Faster R-CNN、YOLOv5、YOLOX、YOLOv7模型的试验对比结果。其中，YOLOv7-tea模型的精确率达到了90.32%，召回率为73.68%，平均精度为80.17%。与原始YOLOv7模型相比，分别提升了2.87个百分点、6.91个百分点和8.69个百分点。说明YOLOv7-tea模型在减少漏检和误检方面取得了显著进步，大幅提高了模型的识别性能。

3）通过模型对不同角度和不同距离的茶叶嫩芽图像进行识别检测效果的测试，确定了YOLOv7-tea模型的识别置信度最高，在茶叶嫩芽识别任务上表现更为出色，由此可见，YOLOv7-tea模型不仅提高了识别精度，而且为茶叶产量预估和茶叶智能化采摘提供了重要的参考。

参　考　文　献

[1] 江用文，陈霄雄，朱建淼，等. 中国茶产业2020年发展规模分析[J]. 茶叶科学，2011，31(3)：273－282.
JIANG Yongwen，CHEN Xiaoxiong，ZHU Jianmiao，et al. Analysis on development scale of Chinese tea industry in 2020 [J]. Journal of Tea Science，2011，31(3)：273－282.

[2] TIAN L，HE Y. Study on the cross-cultural marketing strategies of Chinese tea enterprises [J]. Academic Journal of Business and Management，2022，4(2)：38－41.

[3] 李宗洙. 我国茶产业贸易发展研究：模式、问题及对策建议[J]. 商业经济研究，2021(24)：186－189.

[4] 韩余，宋志禹，陈巧敏. 4CJ－1200F智能采茶机设计与试验[J]. 智能化农业装备学报(中英文)，2022，3(1)：1－6.
HAN Yu，SONG Zhiyu，CHEN Qiaomin. Design and experiment of 4CJ－1200F intelligent tea plucking machine [J]. Journal of Intelligent Agricultural Mechanization，2022，3(1)：1－6.

[5] 张浩，陈勇，汪巍，等. 基于主动计算机视觉的茶叶采摘定位技术[J]. 农业机械学报，2014，45(9)：61－65.
ZHANG Hao，CHEN Yong，WANG Wei，et al. Positioning method for tea picking using active computer vision [J]. Transactions of the Chinese Society for Agricultural Machinery，2014，45(9)：61－65.

[6] 许宝阳，高延峰. 基于Faster-RCNN深度学习的茶叶嫩芽多维度识别及其性能分析[J]. 农业装备与车辆工程，2023，61(2)：19－24.
XU Baoyang，GAO Yanfeng. Tea-buds multi-dimensional recognition with Faster-RCNN deep learning method and its performance analysis [J]. Agricultural Equipment & Vehicle Engineering，2023，61(2)：19－24.

[7] SHAO P D，WU M H，WANG X W，et al. Research on the tea bud recognition based on improved k-means algorithm [C]// MATEC Web of Conferences. EDP Sciences，2018，232：03050.

[8] YAN C，CHEN Z，LI Z，et al. Tea sprout picking point identification based on improved DeepLabV3+ [J]. Agriculture，2022，12(10)：1594.

[9] ZHANG F Y，SUN H W，XIE S，et al. A tea bud segmentation，detection and picking point localization based on the MDY7-3PTB model [J]. Frontiers in Plant Science，2023：1199473－1199473.

[10] CHENG Y F，LI Y，ZHANG R T. Locating tea bud keypoints by keypoint detection method based on convolutional neural network [J]. Sustainability，2023，15(8)：6898.

[11] LI J，LI J H，ZHAO X，et al. Lightweight detection networks for tea bud on complex agricultural environment via improved YOLO v4 [J]. Computers and Electronics in Agriculture，2023，211(C)：107955.

[12] 龙樟，姜倩，王健，等. 茶叶嫩芽视觉识别与采摘点定位方法研究[J]. 传感器与微系统，2022，41(2)：39－41，45.
LONG Zhang，JIANG Qian，WANG Jian，et al. Research on method of tea flushes vision recognition and picking point localization [J]. Transducer and Microsystem Technologies，2022，41(2)：39－41，45.

[13] 杨福增，杨亮亮，田艳娜，等. 基于颜色和形状特征的茶叶嫩芽识别方法[J]. 农业机械学报，2009，40(S1)：119－123.
YANG Fuzeng，YANG Liangliang，TIAN Yanna，et al. Recognition of the tea sprout based on color and shape features [J]. Transactions of the Chinese Society for Agricultural Machinery，2009，40(S1)：119－123.

[14] 王梦妮，顾寄南，王化佳，等. 基于改进YOLOv5s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 YOLOv5s model [J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(12): 150－157.

[15] 俞龙, 黄楚斌, 唐劲驰, 等. 基于YOLOX改进模型的茶叶嫩芽识别方法[J]. 广东农业科学, 2022, 49(7): 49－56.
YU Long, HUANG Chubin, TANG Jinchi, et al. Tea bud recognition method based on improved YOLOX model [J]. Guangdong Agricultural Sciences, 2022, 49(7): 49－56.

[16] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]// Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2023: 7464－7475.

[17] 郝紫霄, 王琦. 基于YOLO-v7的无人机航拍图像小目标检测改进算法[J]. 软件导刊, 2024(1): 167－172.
HAO Zixiao, WANG Qi. Enhanced algorithm for small target detection in UAV aerial images based on YOLO-v7 [J]. Software Guide, 2024(1): 167－172.

[18] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]// Proceedings of the European conference on computer vision (ECCV). 2018: 3－19.

[19] 胡和平, 吴明晖, 洪孔林, 等. 基于改进YOLOv5s的茶叶嫩芽分级识别方法[J]. 江西农业大学学报, 2023, 45(5): 1261－1272.
HU Heping, WU Minghui, HONG Konglin, et al. Classification and recognition method for tea buds based on improved YOLOv5s [J]. Acta Agriculturae Universitatis Jiangxiensis, 2023, 45(5): 1261－1272.

[20] SUNKARA R, LUO T. No more strided convolutions or pooling: A new CNN building block for low-resolution images and small objects [C]// Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Cham: Springer Nature Switzerland, 2022: 443－459.

[21] ZHANG Y F, REN W, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression [J]. Neurocomputing, 2022, 506: 146－157.

Study of tea buds recognition and detection based on improved YOLOv7 model

WEI Tangwei, ZHANG Jincheng, WANG Jing, ZHOU Qingyan*

(*College of Information and Artificial Intelligence, Anhui Agricultural University, Hefei* 230036, *China*)

Abstract: To effectively identify tea buds in complex environments and improve the precision of intelligent harvesting while minimizing damage to tea trees, this study addresses the issues of low detection accuracy and poor robustness exhibited by traditional target detection algorithms in tea gardens, and proposes YOLOv7-tea model for tea bud identification and detection based on an improved YOLOv7, so as to achieve rapid recognition and detection of tea buds. First, tea bud images were collected and annotated, and data augmentation was performed to construct a tea bud dataset. Next, the CBAM attention mechanism module was introduced into three feature extraction layers of the YOLOv7 backbone network to enhance the model's feature extraction capability; the SPD-Conv module was used to replace the SConv module in the neck network's downsampling module to reduce the loss of small object features; and the EIoU loss function was employed to optimize box regression, thereby improving the accuracy of the predicted boxes. Finally, a comparative experiment was conducted between other target detection models and the YOLOv7-tea model using the tea bud image dataset as a sample, and the recognition effect of tea buds shot at different distances and angles was tested. The experimental results show that the YOLOv7-tea network model outperforms the YOLOv7 model in terms of precision (P), recall (R), and mean average precision (mAP) by 2.87, 6.91, and 8.69 percentage points, respectively. Additionally, the model has a faster detection speed and exhibits higher confidence scores in the recognition and detection of tea buds in complex backgrounds. The YOLOv7-tea model constructed in this study demonstrates better recognition performance for small-sized tea leaf buds, reducing instances of missed detection and false alarms. It exhibits good robustness and real-time performance, offering valuable insights for estimating tea yield and implementing intelligent harvesting.

Keywords: tea buds; object detection; CBAM attention mechanism; automated picking; YOLOv7