DOI: 10.12398/j.issn.2096-7217.2025.01.001

基于改进YOLOv5s的轻量化鲜茶叶识别方法

吴擎[1,2]，韦润轩[1]，周乐[1]，杨浩[1]，刘婉茹[1]，徐红梅[1,2*]

（1. 华中农业大学工学院，湖北武汉，430070；2. 农业农村部长江中下游农业装备重点实验室，湖北武汉，430070）

摘要：茶芽分类识别是名优茶生产中十分重要的环节。针对目前茶芽识别算法模型尺寸大、计算量大且无法区分采摘形态等问题，本研究以YOLOv5s为基线模型，提出了一种改进的鲜茶叶识别模型（YOLOv5s-SPCS）。首先，收集实验室环境和自然环境下的鲜茶叶图像制作鲜茶叶数据集。其次，基于ShuffleNetV2思想构建Shuffle Block模块替换主干网络中的卷积模块，在减少模型参数量和计算量的同时提高特征提取速度。然后，在颈部网络引入部分卷积结构PConv和无参数注意力机制SimAM构建C3-PCS模块替换原C3结构，减少模型计算冗余和内存访问，提高识别精度。最后，采用SIoU作为边界框损失函数，提高预测框收敛速度和收敛精度。试验结果表明，YOLOv5s-SPCS模型参数量、计算量和权重文件大小分别为YOLOv5s模型的14%、14%和16%，对鲜茶叶识别准确率为81.8%、平均精度均值为82.4%，相较于原始模型，准确率提升了2.7个百分点，平均精度均值保持不变。此外，改进后的YOLOv5s-SPCS模型整体性能优于当前常用的Faster R-CNN、SSD、YOLOv3、YOLOv4等目标检测模型。本研究可为鲜茶叶识别分类及后续移动端部署提供有效的技术支持。

关键词：深度学习；YOLOv5；轻量化；鲜茶叶；目标检测

中图分类号：S24；TP391.4　　**文献标识码：**A　　**文章编号：**2096-7217（2025）01-0001-14

吴擎，韦润轩，周乐，杨浩，刘婉茹，徐红梅. 基于改进YOLOv5s的轻量化鲜茶叶识别方法[J]. 智能化农业装备学报(中英文)，2025，6(1)：1－14

WU Qing，WEI Runxuan，ZHOU Le，YANG Hao，LIU Wanru，XU Hongmei. Lightweight fresh tea leaf recognition method based on improved YOLOv5s [J]. Journal of Intelligent Agricultural Mechanization，2025，6(1)：1－14

0 引言

茶是世界三大饮料之一，是我国重要的经济作物之一。2022年，我国茶园面积达3 330.3 khm^2，全国干毛茶总产量3 181.0 kt，全国干毛茶总产值达3 180.68亿元，传统六大茶类的生产结构均衡发展[1]，茶产业已成为部分地区的支柱产业。我国幅员辽阔，各茶叶产区气候有所差异，因此形成了诸多具有地域特色的茶叶品牌，如黄山毛峰、安吉白片、敬亭绿雪等，众多的茶叶种类给茶叶生产加工的标准化和自动化进程带来了诸多阻力。

茶叶采摘是茶叶生产加工中重要的一环，采摘时间和茶芽完整性直接决定了成品茶的品质。为了保证茶芽质量，提高采摘效率，降低人工成本，各种采茶机械应运而生[2-5]，但现有采摘机械嫩芽和老叶的“无差别采摘”，使得采摘的茶叶完整度低、质量得不到保证，只能作为廉价大宗茶的生产原料，无法用于名优茶制作[6]。名优茶作为一种高营养、高品质、高经济价值的茶叶，近年来产值占比逐年提高，其采摘要求更高，需要对茶芽进行分级，一般分为无叶单芽、单叶芽、双叶芽3类，如图1所示。

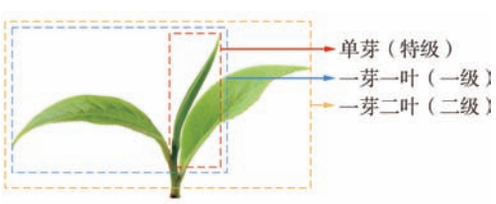


图1 茶芽的不同采摘形态

Figure 1 Different picking patterns of tea buds

现阶段亟须面向名优茶的自动采摘技术，而实现鲜茶叶机械化、自动化采摘的前提和关键便是对茶芽的识别与检测。近年来，茶芽识别检测方法主要包括传统图像处理方法和深度学习方法[7]。黄印[8]利用鲜茶叶相对几何特征与纹理特征基于SVM构建分类器，并基于距离矩阵相似度得到形状特征分类结果，最后

收稿日期：2024-05-14　　修回日期：2024-08-19

基金项目：湖北省高等学校优秀中青年科技创新团队计划项目（T201934）

第一作者：吴擎，女，1988年生，湖北孝感人，博士，副教授；研究方向为深度学习及其应用。E-mail：wuqing@mail.hzau.edu.cn

*通信作者：徐红梅，女，1979年生，湖北随州人，博士，教授；研究方向为智慧农业。E-mail：xhm790912@163.com

利用KNN对两种方法进行结果融合,提出了一种基于多特征与多分类器的鲜茶叶分类方法;范婷婷[9]应用灰度共生矩阵(gray-level co-occurrence matrix, GLCM)从高光谱图像提取纹理特征,基于纹理特征和融合光谱纹理特征分别建立SVM和ELM模型;陈忠辉[10]依据分割后茶叶嫩芽图像的颜色特征,使用RGB颜色分离获取二值化图像,对最大连通区域进行提取,去除背景噪声的干扰,提出了一种基于Shi-Tomasi算法的嫩芽采摘点识别方法。以上方法虽然可以成功识别目标,但在特征选择阶段极依赖人工经验设计,复杂的调参过程耗时耗力[11]。

随着深度学习的快速发展与广泛应用,越来越多的目标检测算法相继问世,其中YOLO系列应用最为广泛[12]。黄家才等[13]对YOLOv4算法进行改进,将主干网络替换为GhostNet,在颈部网络引入Ghost卷积,提出了一种基于Compact-YOLO v4算法的茶叶嫩芽移动端识别方法,并将其部署于PRO-RK3568-B移动端开发板;ZHU等[14]基于3D点云技术和YOLOv5网络实现茶芽定位检测,通过高效通道注意网络(ECANet)模块和双向特征金字塔网络(BiFPN)在获取感兴趣区域的3D点云后,利用DBSCAN聚类算法提取特征获取茶芽采摘点三维坐标,最后利用主成分分析法将最小外接长方体整合在3D点云上,求得三维坐标。GUI等[15]通过引入Ghost Conv减少计算量和模型大小,加入BAM注意力机制,抑制无效信息,再通过加权特征融合方法,提出了一种基于YOLOv5模型的轻量型茶芽识别检测模型。王梦妮等[16]基于YOLOv5s模型,通过引入空洞空间卷积池化金字塔结构ASPP、双向特征金字塔网络BiFPN并在C3模块中添加卷积注意力模块CBAM,提高了模型关注小目标特征的能力和特征融合的效率。然而,现有的茶芽识别算法仍存在模型尺寸大、计算量大、检测速度慢等问题,未能实现识别效率与精度的有效平衡。

综上所述,为提高鲜茶叶目标识别与检测精度和效率,本研究采用YOLOv5s网络作为基线模型,通过引入Shuffle Block轻量化模块、PConv部分卷积结构、SimAM无参数注意力机制和SIoU边界框损失函数提出了一种新的鲜茶叶分类识别模型,在保证性能的同时实现轻量化设计,拟为茶叶识别检测模型搭建和部署提供参考。

1 材料与数据

1.1 图像采集

为实现多场景鲜茶叶识别,本研究选用实验室环境和自然环境下的鲜茶叶图像构建数据集,采集地点为湖北省武汉市华中农业大学南湖教学茶厂(北纬30°48′,东经114°37′),采集对象为福鼎大白,植株高大,树枝半开张,主干较明显,分支较密,叶片呈上斜状生,叶色绿,叶面隆起,有光泽,叶缘平,叶身平,叶尖钝尖。

实验室环境下的鲜茶叶图像由图像采集装置采集,其中图像采集装置由载体平台、光源、光源控制器、摄像机、镜头和计算机组成,如图2所示。图像采集于2023年7月,采集设备为工业相机IDS U3-3890 CPREV. 2. 2,镜头选用UH0820-10M,图像分辨率为4 000像素×3 000像素,以JPG格式保存。在采集过程中,由于鲜茶叶尺寸较小,为了更好地观察其形状和减少冗余背景对目标检测造成的影响,选用纯黑色底板作为拍摄背景,使用点光源通过反射间接点亮样本,将鲜茶叶样本直接放置于相机镜头正下方,调节相机焦距和光圈,使样本与相机镜头保持20 cm距离,立刻采集鲜茶叶图像,并存储在计算机中。

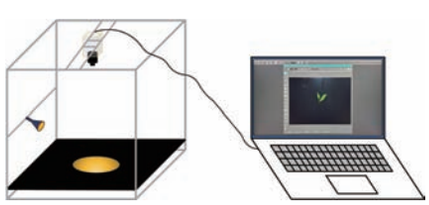


图2 图像采集平台

Figure 2 Image acquisition platform

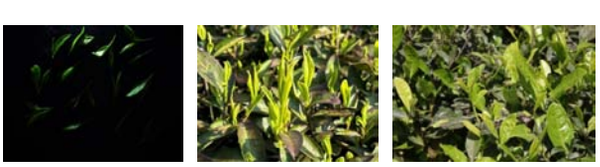
自然环境下的鲜茶叶图像通过手持iPhone 14 Pro进行,拍摄距离维持在30~50 cm之间,图像分辨率为4 032像素×3 024像素,考虑到不同光照条件对图像质量的影响,选择在不同时间段进行图像采集,如图3所示。同时本研究还通过线上收集的方式对自然环境图像进行扩充,最终鲜茶叶数据集组成如图4所示。



(a) 早晨 (b) 中午 (c) 下午

图3 自然环境下的鲜茶叶图像

Figure 3 Images of fresh tea leaves in natural environment



(a) 实验室环境 (b) 自然环境(线下采集) (c) 自然环境(线上收集)

图4 鲜茶叶数据集

Figure 4 Fresh tea leaf dataset

1.2　数据筛选以及图像标注

本研究利用矩形区域标注工具 LabelImg 对鲜茶叶图像进行人工标记，以获得图像中鲜茶叶目标的类别和位置信息，本试验包括“一芽一叶”“一芽二叶”和“单芽”三类主要采摘形态，标记完成的鲜茶叶图像文件以 PASCAL VOC 格式保存在 XML 文件中，带注释的样本与每个图像中的标注框一一对应，每个图像可以包含多个标注框，标注框的多少取决于图像中鲜茶叶的数量。为了满足试验要求，利用 Python 程序将标记好的图像和标签转换为 YOLO 数据格式，该格式将每个图像的标签信息以 (n,x,y,w,h) 的形式保存在 TXT 文件中，其中，n 表示对应类别的序号；x 和 y 分别表示标注框中心点的横、纵坐标；w 和 h 分别表示标注框的相对宽度和高度。鲜茶叶数据集共计 3 197 张图像，按照 8∶2 的比例随机划分为训练集和验证集。

1.3　数据增强

为提升模型的鲁棒性和泛化性，避免过拟合现象，本研究采用多种在线数据增强方式，具体包括：随机裁剪、随机旋转、随机翻转、HSV 颜色空间变换和 Mosaic 增强方式。

其中，随机裁剪指将图像按一定概率进行裁剪；随机旋转指将图像按一定概率、一定角度进行旋转；随机翻转指将图像随机按水平或竖直方向进行镜像处理；HSV 颜色空间变换指对图像色度（Hue）、饱和度（Saturation）和明度（Value）进行随机调整；Mosaic 增强方式指随机使用 4 张图像，通过随机缩放、裁剪、分布进行拼接，得到一张新的图像。各类图像增强效果示意图如图 5 所示。



（a）原图　（b）随机裁剪　（c）随机旋转　（d）随机翻转　（e）HSV 颜色空间变换　（f）Mosaic 增强

图 5　数据增强

Figure 5　Data enhancement

2　鲜茶叶识别算法及改进

2.1　YOLOv5 网络模型

YOLOv5 是 YOLO 系列中推理速度和识别效果兼优模型[17]。其按照模型深度和宽度的参数设置不同分为 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 五个版本，如表 1 所示。按照 NFL 定理[18]，模型性能和模型参数量关系密切，本研究针对三分类鲜茶叶目标，考虑到实时检测和后续移动端部署的需求，采用参数量和计算量相对较少，模型性能也较好的 YOLOv5s 作为基线模型，其结构如图 6 所示。

YOLOv5s 主要由输入层（Input）、主干网络（Backbone）、颈部网络（Neck）和检测网络（Detect）4 部分组成。Input 包含数据增强、图像预处理、自适应锚框计算及图片缩放。Backbone 用于获取多尺度特征图。Neck 通过特征金字塔网络（feature pyramid networks，FPN）[19]自上而下，传递强语义信息，提高模型检测性能，路径聚合网络（path aggregation network，PAN）[20]自下而上，改善模型定位性能。Detect 通过 3 个不同尺度的输出向量，用于预测目标类别概率和位置信息。

表 1　YOLOv5 系列参数对比

Table 1　YOLOv5 series parameter comparison

模型	深度	宽度	参数量 /$\times 10^6$ M	浮点运算量 *FLOPs*/G
YOLOv5n	0.33	0.25	1.76	4.1
YOLOv5s	0.33	0.50	7.02	15.8
YOLOv5m	0.67	0.75	20.86	47.9
YOLOv5l	1.00	1.00	46.15	108.3
YOLOv5x	1.33	1.25	86.23	204.7

2.2　YOLOv5s 网络改进

本研究对 YOLOv5s 模型进行轻量化改进，旨在保证模型性能的同时减少参数量和计算量。固体改进包括：（1）在主干网络中构建 CBRM 模块替换原 6×6 卷积层，基于 ShuffleNetV2 思想构建 Shuffle Block 轻量化模块用于特征提取；（2）在颈部网络中引入部分卷积结构 PConv 和无参数注意力机制 SimAM，构建 PCBS 模块和 C3-PCS 模块替换原卷积模块和 C3 模块；（3）使用 SIoU 作为边界框损失函数。将改进后的模型命名为 YOLOv5s-SPCS，其网络结构如图 7 所示。

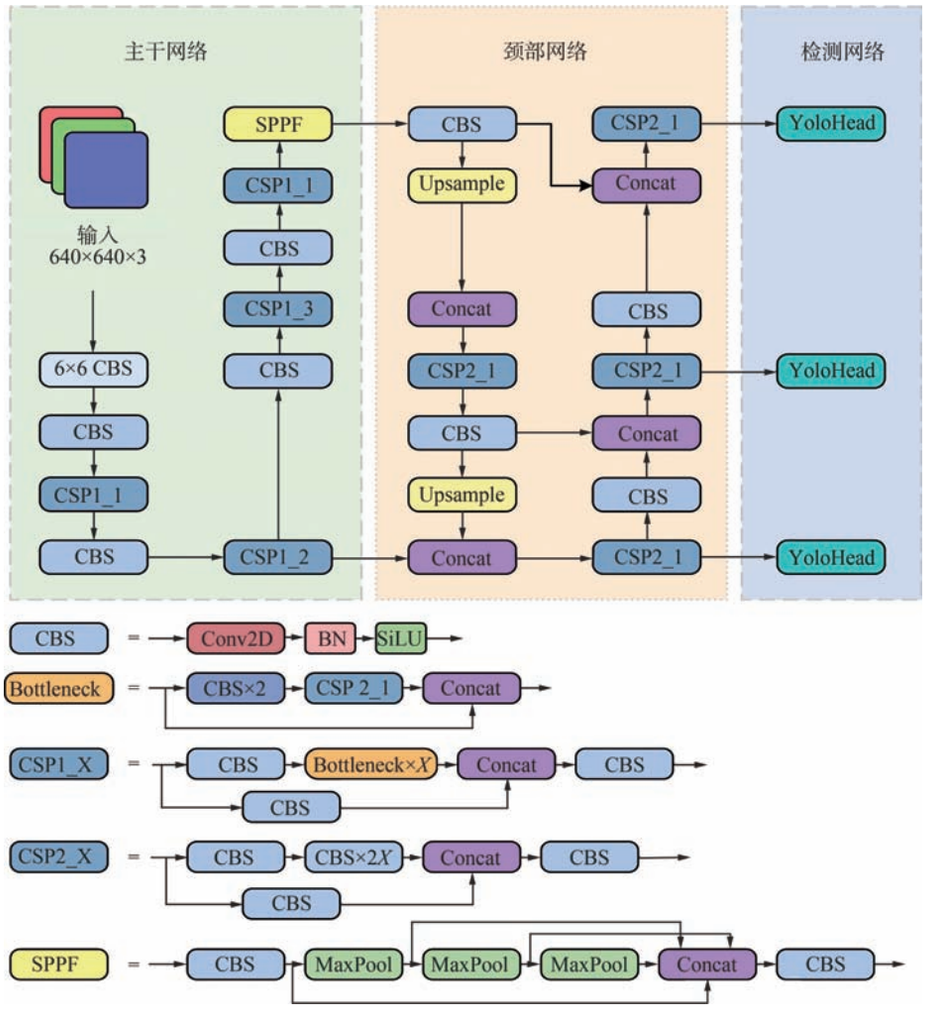


图 6 YOLOv5s 网络结构

Figure 6 The network structure of YOLOv5s

注:Conv2D 为常规卷积操作;BN 为批量归一化操作;SiLU 为激活函数;Bottleneck 为残差结构;Concat 为张量拼接操作;MaxPool 为最大池化操作;SPPF 为空间金字塔池化结构;Upsample 为上采样操作;CBS 模块由 Conv2D、BN 和 SiLU 组成;CSP1_X 为包含 X 个 Bottleneck 的卷积结构;CSP2_X 为包含 $2X$ 个 CBS 的卷积结构。

2.2.1 ShuffleNetV2轻量化网络

ShuffleNetV2[21]是 ShuffleNetV1[22]的改进版本,相较于 V1 版本,其通过引入通道拆分(Channel Split)代替逐点分组卷积(point wise group convolution,GConv),并将通道混洗操作(Channel Shuffle)移至网络末端,设计出速度和准确率兼具的轻量化网络。

ShuffleNetV2 的基本单元如图 8 所示。在每个单元的开始,输入特征通道经过通道拆分被划分为两个分支,右侧分支由 3 个具有相同输入和输出通道的卷积组成,如图 8(a)所示,与 ShuffleNetV1 不同,由于通道拆分已经进行分组操作,所以分支中两个 1×1 卷积操作使用常规卷积而非逐点分组卷积,接着通过张量拼接将两个分支连接,再使用通道混洗实现两个分支之间的信息交互。对于下采样操作(步长 stride 为 2),在图 8(a)的基础上,通过去除通道拆分,并通过在左侧分支上添加一个 3×3 的深度可分离卷积和一个 1×1 的常规卷积用来混合特征,如图 8(b)所示。

本研究基于 ShuffleNetV2 思想设计 Shuffle Block 模块,构建 YOLOv5s-SPCS 的主干网络,并将主干网络第 2、4、6 层提取到的特征图输出,作为后续特征融合层的输入。

2.2.2 PConv部分卷积结构

部分卷积(partial convolution,PConv)是一种新的卷积结构[23],其可以同时减少内存访问和冗余计算,并能够更好地利用设备的计算能力进行特征提取。如图 9 所示,部分卷积的基本原理在于其仅对部分输入通道进行卷积操作,以此减少卷积操作带来的大量计算量,同时通道之间的信息交互也使得单一通道可以成为整体特征图的代表,在特征提取阶段不会影响到整体效果。

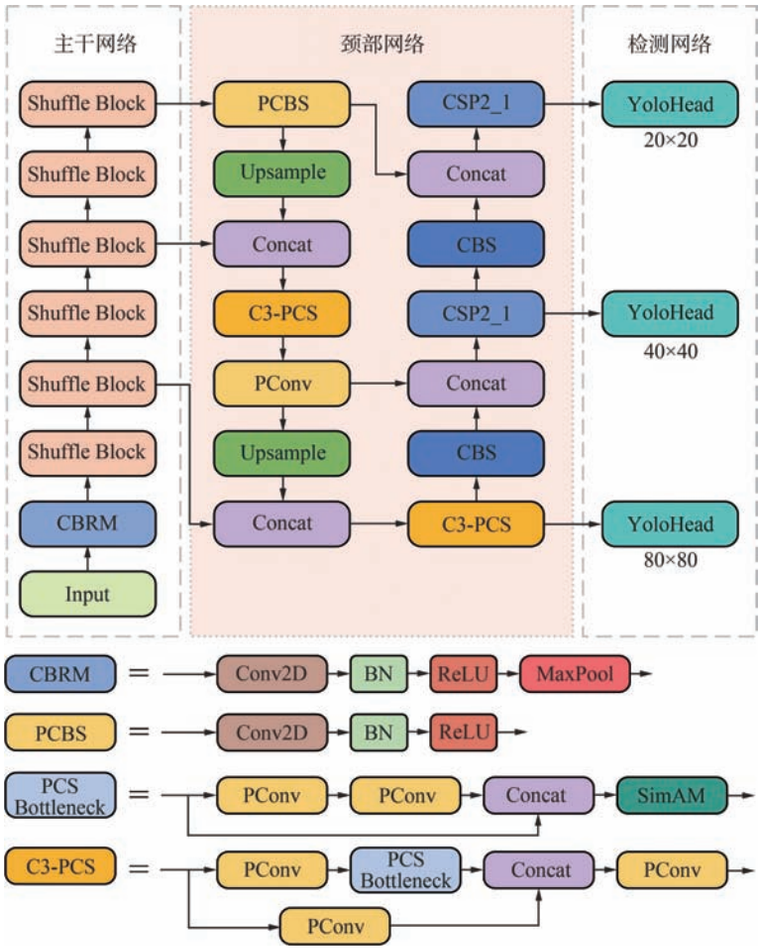
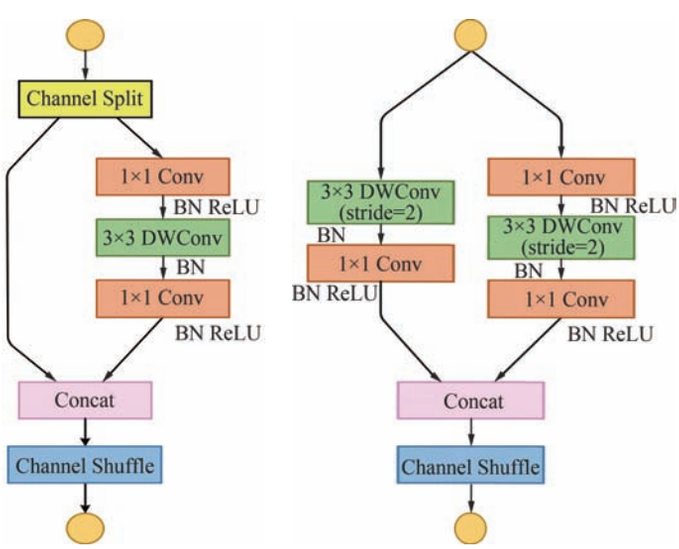


图 7　YOLOv5s-SPCS 网络结构

Figure 7　The network structure of YOLOv5s-SPCS

注：ReLU 为激活函数；PConv 为部分卷积操作；SimAM 为无参数注意力机制模块；PCS Bottleneck 为引入 PConv 和 SimAM 后的残差结构；C3-PCS 为新的 C3 模块；CBRM 模块由 Conv、BN、ReLU 和 MaxPool 组成；PCBS 模块由 PConv、BN 和 SiLU 组成；Shuffle Block 是基本的 ShuffleNet 模块。



（a）常规模块　　（b）下采样模块

图 8　Shuffle Block 轻量化模块结构

Figure 8　Structure of Shuffle Block lightweight module

注：Channel Split 为通道拆分操作；Conv 为常规卷积操作；DWConv（depthwise convolution）为深度可分离卷积；Concat 为张量拼接；Channel Shuffle 为通道重排。

部分卷积的计算量和内存访问成本的数学表达式为

$$FLOPs = H \times W \times K^2 \times C_p^2 \tag{1}$$

$$MAC = H \times W \times 2C_p + K^2 \times C_p^2 \approx H \times W \times 2C_p \tag{2}$$

式中:H——输入特征图的高;

W——输入特征图的宽;

K——卷积核尺寸;

C_p——部分卷积选择进行卷积操作的通道数;

C_{in}——输入特征图通道数;

C_{out}——输出特征图通道数。

一般情况下,$C_p/C_{in}=1/4$ 且 $C_{in}=C_{out}$,此时部分卷积的 $FLOPs$ 为普通卷积的 1/16,MAC 为普通卷积的 1/4,部分卷积的使用能够在不影响特征提取效率的同时减少了计算量。

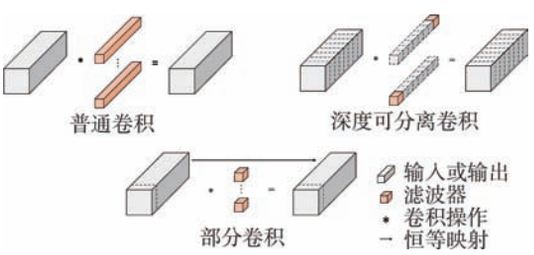


图 9 不同卷积结构

Figure 9 Structure of different convolution

2.2.3 SimAM 无参数注意力机制

SimAM(simple attention module)由 YANG 等人在 2021 年提出[24],其解决了现有注意力机制的两个问题:(1)仅能对通道维度或空间维度进行关注,但对于两个维度同时变化的问题缺乏灵活性;(2)实现注意力权重需要经过一系列复杂计算,会增加模型参数量。其基于神经科学发现所提出的优化能量函数定义每个神经元权重,从而增强重要特征的提取、抑制非重要特征的干扰,计算简单,无须过多的模块调整和额外的参数计算即可实现通道维度和空间维度的注意力关注。现有通道注意力机制、空间注意力机制及 SimAM 的原理图如图 10 所示。优化能量函数 e_t 定义为

$$e_t(w_t, b_t, y, x_i) = \frac{1}{M-1}\sum_{i=1}^{M-1}[-1-(w_i x_i + b_t)]^2 + [1-(w_t t + b_t)]^2 + \lambda w_t^2 \tag{3}$$

式中:t——输入特征 $X \in R^{C*H*W}$ 的目标神经元;

x_i——输入特征 $X \in R^{C*H*W}$ 中与目标神经元同一通道内的其他神经元;

i——空间维度上的索引;

M——通道中的神经元数量,$M=H\times W$;

y——标签值,用于判断目标神经元是否重要;

w_t——变换的权重;

b_t——变换的偏置;

λ——正则化系数。

能量越低,目标神经元 t 与其他神经元的差异越大,相关性越低。

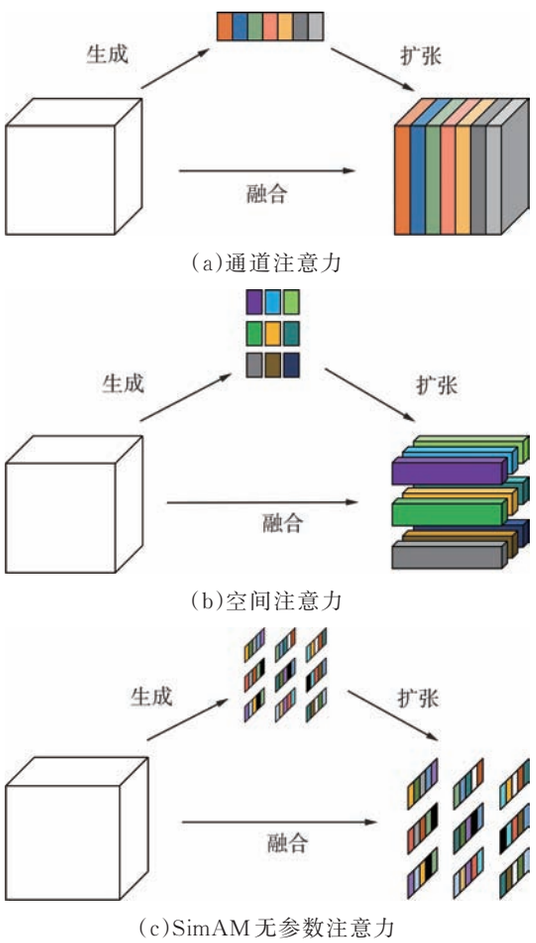


图 10 不同注意力机制对比

Figure 10 Comparison of different attention mechanisms

2.2.4 SIoU 损失函数

平滑交并比(smoothed intersection over union,SIoU)[25]相较于其他边界框损失函数,考虑到期望回归之间向量的角度,重新定义了角度惩罚度量,可以使预测框更快向真实框逼近,提升收敛速度和回归精度。SIoU 损失函数由角度损失 Λ(Anglecost)、距离损失 Δ(Distancecost)、形状损失 Ω(Shapecost)和交并比损失 U 组成,SIoU 参数示意图如图 11 所示。

角度损失

$$\Lambda = 1 - 2 \times \sin^2\left[\arcsin(x) - \frac{\pi}{4}\right] \tag{4}$$

其中:

$$x = \frac{c_h}{\sigma} = \sin(\alpha) \tag{5}$$

距离损失

$$\Delta=\sum_{t=x,y}(1-e^{-\gamma\rho t}) \tag{6}$$

其中：

$$\rho_x=\left(\frac{b_{cx}^{gt}-b_{cx}}{c_w}\right)^2,\rho_y=\left(\frac{b_{cy}^{gt}-b_{cy}}{c_h}\right)^2,\gamma=2-\Lambda \tag{7}$$

形状损失

$$\Omega=\sum_{t=w,h}(1-e^{-\omega_t})^{\theta} \tag{8}$$

其中：

$$\omega_w=\frac{|w-w^{gt}|}{\max(w,w^{gt})},\omega_h=\frac{|h-h^{gt}|}{\max(h,h^{gt})} \tag{9}$$

式中：θ——对形状损失的关注程度，以避免过度关注形状损失降低预测框的移动，$\theta\in(2,6)$；

w——预测框的宽；

h——预测框的高；

w^{gt}——真实框的宽；

h^{gt}——真实框的高。

SIoU 损失的定义为

$$L_{SIoU}=1-U+\frac{\Delta+\Omega}{2} \tag{10}$$

其中：

$$U=\frac{|B\cap B^{GT}|}{|B\cup B^{GT}|} \tag{11}$$

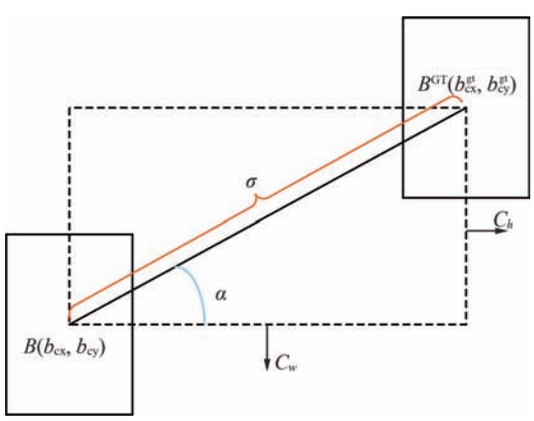


图 11　SIoU 参数示意图

Figure 11　SIoU parameters schematic diagram

注：B 和 B^{GT} 分别为预测框中心点和真实框中心点，(b_{cx},b_{cy}) 和 $(b_{cx}^{gt},b_{cy}^{gt})$ 分别是 B 和 B^{GT} 的坐标。C_w 和 C_h 分别是 B 和 B^{GT} 横纵坐标的差值，α 为 B 和 B^{GT} 连线与水平轴夹角，σ 为 B 和 B^{GT} 连线长度。

3　试验方法及结果

3.1　试验环境与参数设置

本研究试验平台基于 Windows10 的 64 位操作系统，12 th Gen Intel® Core™ i7-12700KF、3.6 GHz 处理器和 NVIDIA GeForce RTX 3070Ti 显卡，显存大小 8 GB，GPU 加速库为 CUDA11.6 和 CUDNN8.2，Python 版本为 3.7.12，深度学习框架为 PyTorch1.12.0。

试验采用随机梯度下降法（stochastic gradient descent，SGD）作为网络的优化器进行优化，为保持模型深层的稳定性，训练采用 warmup 策略，在前 3 轮内采用 0.000 1 的学习率进行预热训练，之后恢复初始学习率，学习率调整方式采用余弦退火学习率衰减策略，初始学习率设置为 0.01，最终学习率设置为 0.000 1，权重衰减系数设置为 0.000 5，动量因子设置为 0.937，共迭代 300 轮。训练时将输入图像分辨率统一调整为 640 像素×640 像素，批次大小设置为 16，所有试验均为从头开始训练，未使用预训练权重文件和迁移学习方法。

3.2　评价指标

本研究选择目标检测模型中常用的评价指标：精确度（precision，P）、召回率（recall，R）、平均精度均值（mean average precision，mAP）、模型参数量（parameters，$Params$）、浮点运算量（floating point of operation，$FLOPs$）、每秒传输帧率（frames per second，FPS）和权重文件大小，对改进后的模型进行综合评价。具体定义如下：

$$P=\frac{TP}{TP+FP} \tag{12}$$

$$R=\frac{TP}{TP+FN} \tag{13}$$

$$AP=\int_0^1(P\times R)\mathrm{d}x \tag{14}$$

$$mAP=\frac{1}{N}\sum_{i=1}^{N}AP(i) \tag{15}$$

式中：TP——实际正类预测为正类的数量，即模型正确检测的目标数目；

FN——实际正类预测为负类的数量，即模型错误检测的目标数目；

FP——实际负类预测为正类的数量，即模型将背景检测为目标的数目。

3.3　消融试验

本研究以 YOLOv5s 作为基线模型，基于 ShuffleNetV2 思想构建 Shuffle Block 轻量化卷积模块用于主干网络，以降低模型参数量和计算量；在颈部网络引入部分卷积结构，更大程度地利用设备计算能力，同时引入无参数注意力机制，在不增加额外参数量的前提下使模型更加重视感兴趣区域；使用 SIoU 作为边界框损失函数，提高预测框的收敛速度。通过消融试验，验证不同模块组合对于网络性能的影响，消融试验结果如表 2 所示。

试验结果表明,改进后的YOLOv5s-SPCS算法较原始模型在各项指标上均有提升。改进后的模型精确率为81.8%,召回率为72.5%,较原始模型分别提升了2.7和0.4个百分点,模型平均精度均值为82.4%,与原始模型保持一致。改进后的模型参数量为0.96 M,浮点运算量为2.2 G,较原始模型分别减少了86.3%和86.1%。改进后的模型所保存的权重文件大小为2.09 MB,较原始模型减少了85%,满足实时检测的要求。当主干网络使用Shuffle Block时,由于其通道划分操作的原理,模型在保持性能的同时大幅减少了参数量和计算量。当颈部网络使用部分卷积替换常规卷积时,由于部分卷积仅对部分通道进行常规卷积处理的原因,在小幅增加参数量和计算量的前提下,获得了更好的识别性能。当SimAM应用于C3模块时,在不增加参数量的前提下,模型特征提取环节加入三维权重,使得识别精度得到提升。综上,改进后的模型在精确度、召回率、*mAP*上均有小幅提升,参数量、浮点运算量和权重文件大小均有大幅减少。

表2 网络消融试验

Table 2 Network ablationtest

Shuffle Block	PConv	SimAM	SIoU	精确度 *P*/%	召回率 *R*/%	平均精度均值 *mAP*/%	参数量 *Params*/×10^6 M	浮点运算量 *FLOPs*/G	权重文件大小/MB
×	×	×	×	79.1	72.1	82.4	7.02	15.8	13.67
√	×	×	×	78.2	69.1	80.4	0.85	1.8	1.85
×	√	×	×	82.3	72.3	82.1	7.20	16	14.03
×	×	√	×	82.1	72.7	82.5	7.29	17.2	14.20
√	√	×	×	79.5	68.2	81.4	0.89	1.9	1.95
×	√	√	×	80.9	73.1	82.6	7.48	17.4	14.56
√	×	√	×	78.1	70.3	80.9	0.91	2.2	1.99
√	√	√	×	81.8	71.9	81.2	0.96	2.2	2.09
√	√	√	√	81.8	72.5	82.4	0.96	2.2	2.09

注:"×"表示不使用此模块;"√"表示使用此模块。

YOLOv5s-SPCS模型训练过程中损失值、精确率、召回率、平均精度均值等性能参数变化曲线如图12所示,对应三分类鲜茶叶目标检测性能如表3所示。

表3 YOLOv5s-SPCS模型三分类目标检测性能

Table 3 Three-classification target detection performance of YOLOv5s-SPCS

类别	精确率 *P*/%	召回率 *R*/%	平均精度均值 *mAP*/%
全类别	81.8	72.5	82.4
一芽一叶	79.5	75.5	85.3
一芽二叶	86.5	79.1	89.4
单芽	79.9	61.2	72.1

YOLOv5s-SPCS模型对应三分类鲜茶叶目标检测性能如表3所示。在精确率、召回率和平均精度均值3个指标中,模型对一芽二叶目标检测性能最优;而对单芽目标检测性能较差。究其原因,一芽二叶目标相较于一芽一叶和单芽两个类别目标尺寸更大,更利于模型识别,而单芽目标相对较小,对模型识别提出挑战。

3.4 不同轻量化方法对比

为验证不同主干网络轻量化改进方法的性能差异,本研究分别基于MobileNetV3[26]、GhostNet[27]和ShuffleNetV2对YOLOv5s的主干网络进行改进,试验结果如表4所示。

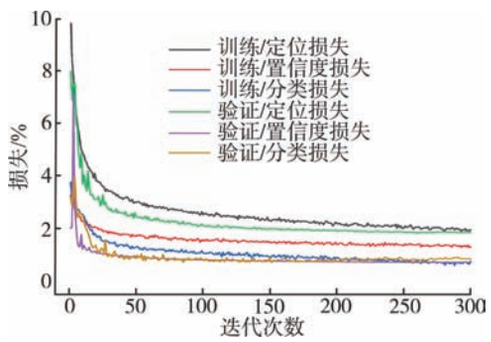
试验结果表明,基于ShuffleNetV2进行改进时,模型的参数量和浮点运算量最低,分别为0.96 M和2.2 G,相较于基于MobileNetV3和GhostNet分别减少了0.42 M、0.1 G和4.12 M、8.4 G。模型大小为2.09 MB,分别较基于MobileNetV3和GhostNet的两个模型减小了0.83 MB和7.94 MB。基于GhostNet构建主干网络时,模型平均精度均值最高,达到83.1%,略高于MobileNetV3的81.9%和ShuffleNetV2的82.4%,但其模型参数量和浮点运算量均远高于另外两组。基于ShuffleNetV2构建主干网络的模型帧率为67.87帧/s,较基于MobileNetV3和GhostNet分别提高了28.83%和77.07%。综上所述,基于ShuffleNetV2对YOLOv5s主干网络进行改进时,能够

在实现轻量化的同时，获得更好的性能表现，同时模型大小尺寸最小，拥有最高的检测速度。

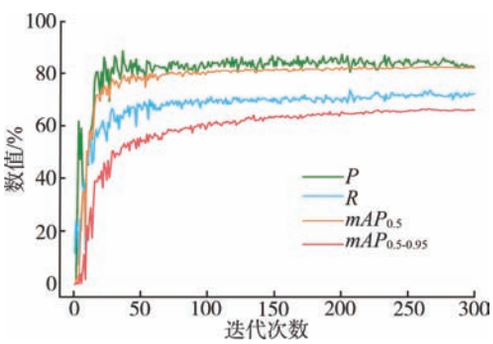
表 4　不同轻量化方法对比试验

Table 4　Comparison experimental of different lightweight approaches

主干网络	平均精度均值 *mAP*/%	参数量 *Params*/×10⁶M	浮点运算量 *FLOPs*/G	模型大小 /MB	帧率 *FPS*/(帧·s⁻¹)
MobileNetV3	81.9	1.38	2.3	2.92	52.68
GhostNet	83.1	5.08	10.6	10.03	38.33
ShuffleNetV2	82.4	0.96	2.2	2.09	67.87



(a) 损失值变化曲线



(b) 精确率、召回率、$mAP_{0.5}$、$mAP_{0.5\text{-}0.95}$变化曲线

图 12　YOLOv5s-SPCS 模型训练过程

Figure 12　Model training process of YOLOv5s-SPCS

3.5　不同损失函数试验对比

目标检测网络的损失由边界框损失、置信度损失和分类损失 3 部分组成。

其中边界框损失主要包括预测框和真实框重叠面积、中心点距离和宽高比 3 部分，原始的 IoU 仅考虑了重叠面积，后续版本均是在其基础上进行改进。GIoU 引入了最小外接矩形和差集概念，解决了边界框不相交的问题[28]；DIoU 引入了中心距离概念，解决了 GIoU 收敛速度较慢的问题[29]；CIoU 充分考虑了 3 个方面，进一步实现了精度的提升；EIoU 进一步优化了 CIoU 中纵横比的定义[30]；SIoU 在 EIoU 基础上增加了角度损失度量，一定程度上提高了收敛速度。为了验证 SIoU 在模型改进过程中的性能提升，本节使用 CIoU、GIoU、DIoU 和 EIoU 进行对比试验，试验结果如表 5 和图 13 所示。

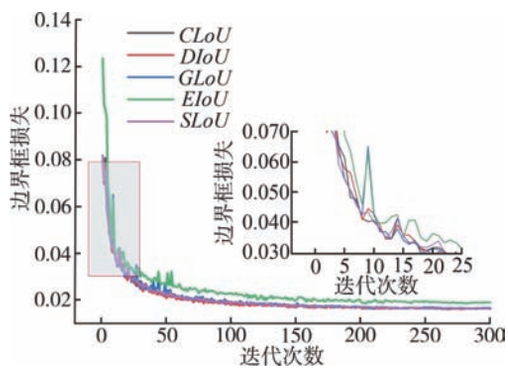


图 13　不同边界框损失函数对比试验

Figure 13　Comparison experimental using different IoU

由表 5 可知，使用 SIoU 作为边界框损失函数时，YOLOv5s-SPCS 模型的平均精度均值最高，为 82.4%，与 EIoU 组相同，分别高于 GIoU、DIoU 和 CIoU 组 0.7、0.1 和 0.9 个百分点；检测时间最短，为 14.7 ms，低于 GIoU、DIoU、EIoU 和 CIoU 组；每秒检测图像数目最多，为 67.87 张。

由图 13 可知，各损失函数对应的边界框损失收敛数值相近，但使用 SIoU 作为边界框损失函数时，损失朝收敛方向下降速度最快，最先实现收敛，该试验验证了 SIoU 的改进效果。

3.6　不同目标检测模型的试验对比

将改进后的模型与其他常用目标检测模型，如 Faster R-CNN[31-32]、SSD-300[33]、YOLOv3[34]、YOLOv4[35]、YOLOv7[36] 以及轻量级目标检测模型 YOLOv3-tiny 和 YOLOv4-tiny 进行对比，试验结果如表 6 所示。试验结果表明，YOLOv5s-SPCS 模型精确率为 81.8%，相较于 YOLOv3、YOLOv4 和 YOLOv5s 分别提升了 0.9、1.6 和 2.7 个百分点，略低于 YOLOv7 83.3% 的精确率，但考虑到 YOLOv7 模型训练过程较长，共花费时间 45.6 h，改进后 YOLOv5s-SPCS 模型训练时长仅为其 10% 左右，实现了轻量化和高性能的平衡。YOLOv5s-SPCS 模型平均精度均值为 82.4%，同改进前保持一致，结合表 2 可知，改进

后模型召回率为72.5%,较原模型提高0.4个百分点。改进后模型平均精度均值较双阶段目标检测模型Faster R-CNN提高了10.2个百分点;较其他单阶段目标检测模型SSD-300、YOLOv3、YOLOv7分别提高了18.7、0.3和0.5个百分点,略低于YOLOv4 82.5的平均精度均值。在参数量和浮点运算量对比中,YOLOv5s-SPCS模型参数量为0.96 M,较Faster R-CNN、SSD-300、YOLOv3、YOLOv4、YOLOv5s和YOLOv7分别减少了99.30%、95.98%、98.44%、98.41%、86.32%和97.42%;浮点运算量为2.2 G,分别减少了99.41%、96.39%、98.58%、98.32%、86.08%和97.91%;YOLOv5s-SPCS模型大小仅为2.09 MB,分别是Faster R-CNN、SSD-300、YOLOv3、YOLOv4、YOLOv5s、YOLOv7模型大小的0.40%、2.28%、1.78%、1.80%、15.29%和2.93%。与轻量级目标检测模型YOLOv3-tiny和YOLOv4-tiny相比,YOLOv5s-SPCS模型精确率和平均精度均值更佳,且在模型参数量、运算量方面更优,分别减少了88.93%和68.73%。

上述对比试验结果表明,本章对YOLOv5s的改进效果显著,在实现了轻量化的同时获得了较高的茶叶目标识别检测精度。

表5 不同边界框损失函数对比试验

Table 5 Comparison experimental using different IoU

损失函数	平均精度均值 mAP/%	检测时间 /ms	帧率 FPS/(帧·s^{-1})
GIoU	81.7	15.2	65.42
DIoU	82.3	15.1	66.30
CIoU	81.5	15.2	66.21
EIoU	82.4	14.8	67.32
SIoU	82.4	14.7	67.87

表6 不同检测模型对比试验

Table 6 Comparison experimental of different detection models

模型	精确率 P/%	平均精度均值 mAP/%	参数量 $Params$/×10^6 M	浮点运算量 $FLOPs$/G	模型大小 /MB
Faster R-CNN	—	72.2	136.73	369.8	521.60
SSD-300	—	63.7	23.88	61.0	91.63
YOLOv3	80.9	82.1	61.51	154.6	117.73
YOLOv3-tiny	81.3	79.4	8.67	12.9	16.63
YOLOv4	80.2	82.5	60.40	130.7	115.80
YOLOv4-tiny	79.6	81.6	3.07	6.4	5.96
YOLOv5s	79.1	82.4	7.02	15.8	13.67
YOLOv7	83.3	81.9	37.21	105.1	71.32
YOLOv5s-SPCS	81.8	82.4	0.96	2.2	2.09

3.7 模型特征可视化

为更直观观察到网络对不同目标的学习情况,将改进后的模型与基线模型训练过程可视化,选取部分结果作为展示,如图14所示。其中,上方为原始模型YOLOv5s的特征提取可视化效果,下方为YOLOv5s-SPCS。图14(b)、(c)、(d)、(e)分别为图像经过模型主干网络第一层卷积结构、颈部网络第一层卷积结构、第二层上采样结构和输出网络前的特征可视化效果,理论上,随着模型深度的增加,所提取到的特征深度也会随之增加,表现在特征图上即为图像更为抽象。

同时,采用Grad-CAM方法(gradient-weighted class activation mapping)[37]进行热力图绘制,可以较为直观地展示模型的学习情况以评价模型的识别性能。Grad-CAM是利用权重文件进行反向传播,对特征图各个通道进行加权激活,获取热力图。模型改进前后热力图如图15所示。改进前后模型对目标检测目标颜色相近,证明了本研究采用的轻量化改进方法能够在不影响检测精度的同时实现。

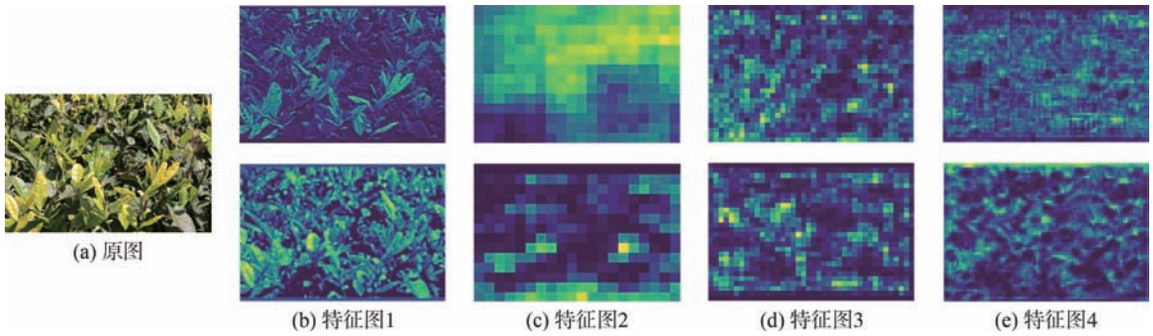


图 14　模型特征可视化

Figure 14　Model feature visualization

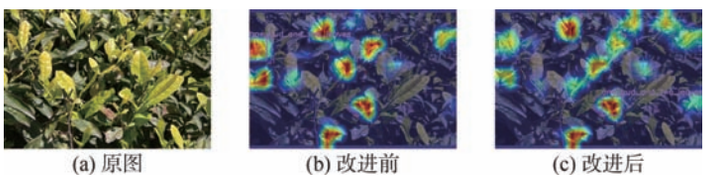


图 15　模型热力图可视化

Figure 15　Model heat map visualization

3.8　模型检测效果

基于改进前后的 YOLOv5s 模型进行测试，目标检测效果如图 16 所示，其中红框表示漏检或误检，蓝框表示小目标样本检测。检测结果表明，改进后的 YOLOv5s-SPCS 模型在该图像检测中，较基线模型生成了更多的检测框，虽然存在误检现象，但没有出现漏检，且能够正确地检测出图像中的小目标样本，证明了改进后模型检测鲜茶叶目标的能力有所提升。



图 16　改进前后目标识别结果对比

Figure 16　Comparison of tea leaf recognition results before and after improvement

4　结论

本研究提出了一种基于改进 YOLOv5s 模型的轻量化鲜茶叶识别方法，可以快速、准确地实现实验室环境和自然环境下鲜茶叶目标的分类识别，主要结论如下。

1）针对现有茶芽目标检测模型尺寸大、计算量大、检测速度慢等问题，以 YOLOv5s 作为基线模型，基于轻量级卷积神经网络 ShuffleNetV2 思想构建 ShuffleBlock 轻量化模块，替换原有卷积结果，大幅减少模型参数量和计算量，提高特征提取效率。

2）针对主干网络轻量化改进后，模型性能下降的问题，在颈部网络中引入部分卷积结构 PConv 和无参数注意力机制 SimAM，在小幅增加计算量的同时，提高模型的特征融合能力，注意力机制的加入使得模型更加注重重点特征。

3）选用 SIoU 作为边界框损失函数，加速预测框收敛速度的同时提高精确率，改进后的 YOLOv5s 模型精确率达到 81.8%，平均精度均值达到 82.4%，优于现有目标检测模型 SSD、Faster R-CNN、YOLOv3、YOLOv4 和 YOLOv5s 等。改进后的 YOLOv5s 模型参数量为 0.96 M，浮点运算量 2.2 G，权重文件大小 2.09 MB，在实现轻量化的同时具有优异的识别性能。

参　考　文　献

[1] 梅宇，张朔．2022年中国茶叶生产与内销形势分析[J]．中国茶叶，2023，45(4)：25－30.
MEI Yu，ZHANG Shuo．Analysis of China′s tea production and domestic sales in 2022 [J]．China Tea，2023，45(4)：25－30.

[2] 高震宇，王安，刘勇，等．基于卷积神经网络的鲜茶叶智能分选系统研究[J]．农业机械学报，2017，48(7)：53－58.
GAO Zhenyu，WANG An，LIU Yong，et al．Intelligent fresh-tea-leaves sorting system research based on convolution neural network [J]．Transactions of the Chinese Society for Agricultural Machinery，2017，48(7)：53－58.

[3] 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.

[4] 胡程喜, 谭立新, 王文胤, 等. 基于改进DeepLabV3＋的轻量化茶叶嫩芽采摘点识别模型[J]. 智慧农业(中英文), 2024, 6(5): 119－127.
HU Chengxi, TAN Lixin, WANG Wenyin, et al. Lightweight tea shoot picking point recognition model based on improved DeepLabV3＋ [J]. Smart Agriculture, 2024, 6(5): 119－127.

[5] 张兰兰, 董迹芬, 唐萌, 等. 名优茶机采鲜叶分级技术研究[J]. 浙江大学学报(农业与生命科学版), 2012, 38(5): 593－598.
ZHANG Lanlan, DONG Jifen, TANG Meng, et al. Classification technology of machine-plucking high quality tea [J]. Journal of Zhejiang University (Agriculture and Life Sciences), 2012, 38(5): 593－598.

[6] 唐小林, 李文萃, 范起业. 机采茶鲜叶分类分级技术及相关设备研究进展[J]. 中国茶叶加工, 2015(2): 5－8.
TANG Xiaolin, LI Wencui, FAN Qiye. Research on the classification technology and the grading equipment of machine-plucking fresh leaves [J]. China Tea Processing, 2015(2): 5－8.

[7] 孙道宗, 张振宇, 陈俊聪, 等. 一种基于深度学习的端到端生菜无损鲜重估测模型的建立[J]. 南京农业大学学报, 2024, 47(6): 1212－1220.
SUN Daozong, ZHANG Zhenyu, CHEN Juncong, et al. A model for end-to-end non-destructive fresh weight estimation of lettuce based on deep learning [J]. Journal of Nanjing Agricultural University, 2024, 47(6): 1212－1220.

[8] 黄印. 基于机器学习的鲜茶叶分类研究[D]. 武汉: 中南民族大学, 2021.
HUANG Yin. Research on classification of fresh tea based on machine learning [D]. Wuhan: South-Central Minzu University, 2021.

[9] 范婷婷. 基于高光谱成像技术的茶叶无损检测方法研究[D]. 天津: 河北工业大学, 2021.
FAN Tingting. Research on nondestructive detection method of tea based on hyperspectral imaging technology [D]. Tianjin: Hebei University of Technology, 2021.

[10] 陈忠辉. 基于计算机视觉的茶叶嫩芽识别方法研究[D]. 贵阳: 贵州大学, 2022.
CHEN Zhonghui. Research on the method of tea sprout recognition based on computer vision [D]. Guiyang: Guizhou University, 2022.

[11] 邢洁洁, 谢定进, 杨然兵, 等. 基于YOLOv5s的农田垃圾轻量化检测方法[J]. 农业工程学报, 2022, 38(19): 153－161.
XING Jiejie, XIE Dingjin, YANG Ranbing, et al. Lightweight detection method for farmland waste based on YOLOv5s [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(19): 153－161.

[12] SHANG Y Y, XU X S, JIAO Y T, et al. Using lightweight deep learning algorithm for real-time detection of apple flowers in natural environments [J]. Computers and Electronics in Agriculture, 2023, 207: 107765.

[13] 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 Jiacai, 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.

[14] ZHU L X, ZHANG Z H, LIN G C, et al. Detection and localization of tea bud based on improved YOLOv5s and 3D point cloud processing [J]. Agronomy, 2023, 13(9): 2412.

[15] GUI Z Y, CHEN J N, LI Y, et al. A lightweight tea bud detection model based on Yolov5 [J]. Computers and Electronics in Agriculture, 2023, 205: 107636.

[16] 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 Jinan, 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.

[17] 白强, 高荣华, 赵春江, 等. 基于改进YOLOV5s网络的奶牛多尺度行为识别方法[J]. 农业工程学报, 2022, 38(12): 163－172.
BAI Qiang, GAO Ronghua, ZHAO Chunjiang, et al. Multi-scale behavior recognition method for dairy cows based on improved YOLOV5s network [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(12): 163－172.

[18] WOLPERT D H, MACREADY W G. No free lunch theorems for optimization [J]. IEEE Transactions on Evolutionary Computation, 1997, 1(1): 67－82.

[19] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]// 2017 IEEE

Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey: IEEE, 2017: 936－944.

[20] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation [C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, New Jersey: IEEE, 2018: 8759－8768.

[21] MA N N, ZHANG X Y, ZHENG H T, et al. ShufflenetV2: Practical guidelines for efficient CNN architecture design [C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 116－131.

[22] ZHANG X Y, ZHOU X Y, LIN M X, et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 6848－6856.

[23] CHEN J R, KAO S H, HE H, et al. Run, Don't walk: Chasing higher FLOPS for faster neural networks [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 12021－12031.

[24] YANG L X, ZHANG R Y, LI L D, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks [C]// International Conference on Machine Learning. PMLR, 2021: 11863－11874.

[25] GEVORGYAN Z. SIoU loss: More powerful learning for bounding box regression [J]. arXiv preprint arXiv: 2205.12740, 2022.

[26] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]// 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, New Jersey: IEEE, 2019: 1314－1324.

[27] HAN K, WANG Y H, TIAN Q, et al. GhostNet: More features from cheap operations [C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey: IEEE, 2020: 1577－1586.

[28] REZATOFIGHI H, TSOI N, GWAK J, 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, New Jersey: IEEE, 2019: 658－666.

[29] ZHENG Z H, WANG P, REN D W, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation [J]. IEEE Transactions on Cybernetics, 2022, 52(8): 8574－8586.

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

[31] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39 (6): 1137－1149.

[32] 周华茂，王婧，殷华，等. 基于改进 Mask R-CNN 模型的秀珍菇表型参数自动测量方法[J]. 智慧农业(中英文)，2023，5(4)：117－126.
ZHOU Huamao, WANG Jing, YIN Hua, et al. Phenotype analysis of Pleurotus geesteranus based on improved Mask R-CNN [J]. Smart Agriculture, 2023, 5(4): 117－126.

[33] 郭小燕，于帅卿. 一种轻量级 YOLOv5S 农作物虫害目标检测模型[J]. 南京农业大学学报，2024，47(5)：1009－1018.
GUO Xiaoyan, YU Shuaiqing. A lightweight YOLOv5S crop pest target detection model[J]. Journal of Nanjing Agricultural University, 2024, 47(5): 1009－1018.

[34] REDMON J, FARHADI A. Yolov3: An incremental improvement [J]. arXiv preprint arXiv: 1804.02767, 2018.

[35] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal speed and accuracy of object detection [J]. arXiv preprint arXiv: 2004.10934, 2020.

[36] 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]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey: IEEE, 2023: 7464－7475.

[37] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad-CAM: Visual explanations from deep networks via gradient-based localization [C]// Proceedings of the IEEE International Conference on Computer Vision. 2017: 618－626.

Lightweight fresh tea leaf recognition method based on improved YOLOv5s

WU Qing[1, 2], WEI Runxuan[1], ZHOU Le[1], YANG Hao[1], LIU Wanru[1], XU Hongmei[1, 2*]

(1. *College of Engineering, Huazhong Agricultural University, Wuhan* 430070, *China*; 2. *Key Laboratory of Agricultural Equipment in Mid-Lower Yangtze River Ministry of Agriculture and Rural Affairs, Wuhan* 430070, *China*)

Abstract: The classification and recognition of tea buds represents a crucial aspect of renowned tea production. In view of the problems of large model size, large computational complexity and inability to distinguish the picking morphology of the current tea bud recognition algorithm, this study proposes an enhanced fresh tea leaf recognition model (YOLOv5s-SPCS) based on YOLOv5s as the foundational model. Firstly, images of fresh tea leaves were collected in both laboratory and natural environments to create a dataset of fresh tea leaves. This was done through offline and online collection of images in multiple scenarios, with the resulting images divided into a training set and a test set. Secondly, the Shuffle Block module was constructed based on the ShuffleNetV2 idea for replacing the convolution module in YOLOv5s backbone network, which reduced the number of model parameters and the amount of computation while increasing the speed of feature extraction. Subsequently, the Partial Convolution structure, PConv and SimAM were incorporated into the neck network to construct the C3-PCS module, replacing the original C3 structure which further reduced the model computational redundancy and memory access, while improving the recognition accuracy with a minimal increase in the number of parameters. Finally, the SIoU bounding box loss function was employed to enhance the convergence velocity and precision of the prediction frame. In addition to accelerating the convergence of the model prediction frame regression, the use of this loss function also generates more accurately positioned prediction frames. The experimental results demonstrate that the enhanced YOLOv5s-SPCS model exhibits 14%, 14% and 16% of the YOLOv5s model in terms of the number of parameters, computational volume and weight file. The size of the model is, respectively, with an accuracy of 81.8% and a mean average precision (*mAP*) of 82.4% for the fresh tea image recognition, which is 2.7% more accurate than the original model. The accuracy was enhanced by 2.7 percentage points, while the mean average precision of *mAP* remained unaltered. Furthermore, the overall performance of the enhanced YOLOv5s-SPCS model is superior to that of the prevailing target detection models, including Faster R-CNN, SSD, YOLOv3, and YOLOv4. This study offers a valuable technical foundation for fresh tea leaves recognition classification and subsequent mobile deployment.

Keywords: deep learning; YOLOv5s; lightweight; fresh tea; target detection

Foundation items: The Outstanding Young and Middle-aged Science and Technology Innovation Team Program of Colleges and Universities of Hubei Province (T201934)

First Author: WU Qing, E-mail: wuqing@mail.hzau.edu.cn

* **Correspondence Author:** XU Hongmei, E-mail: xhm790912@163.com