基于改进Yolov5s的玉米喇叭口检测方法

李文强，宋雪航，宋忠强，吕仕力，户鞘楠，王亚玲
（河南农业大学理学院，河南 郑州 450002）

摘要：【目的】提出一种改进的YOLOv5s玉米喇叭口检测算法模型，旨在实现模型网络结构轻量化的同时，提高检测准确率，以提升玉米精准喷药的效率和减少环境污染。【方法】在YOLOv5s网络结构基础上，首先，使用FasterNet结构替换YOLOv5s的C3结构实现模型的轻量化；随后，研究了挤压-激励（Squeeze-and-Excitation，SE）和卷积注意力模块（Convolutional Block Attention Module，CBAM）两种注意力机制分别位于主干网络不同位置时对模型性能的影响；最后，对比分析YOLOv5s 中CIoU，EIoU和WIoU损失函数对模型精度的影响。【结果】注意力机制对比实验表明，CBAM和SE注意力机制在特定位置提升了模型检测精度，最高分别为89.1%和88.8%，但两者同时添加至主干网络时，精度却下降至84.6%。相较于CIoU，EIoU，使用WIoU损失函数对模型性能提升更加有效，其精度和召回率能够达到87.5%和89.7%。经过改进后的模型在检测数据集上的*mAP*@0.5达到了92.3%，较改进前的YOLOv5s（87.8%）提升了4.5%，参数量下降18%，运算量减少21%，并保持了较快的检测速度。【结论】改进后的算法在玉米喇叭口检测任务中表现优异，能够在有限计算资源条件下实现高效、精准的喷药决策支持。该研究为农业精准喷药技术提供了创新的解决方案，也为未来智能农业设备的开发奠定了技术基础。

关键词：玉米喇叭口；目标检测；YOLOv5s算法；精准喷洒
中图分类号：S126　文献标志码：A

Detection method of corn bell mouth based on improved Yolov5s

LI Wenqiang, SONG Xuehang, SONG Zhongqiang, LÜ Shili, HU Qiaonan, WANG Yaling
(College of Science, Henan Agricultural University, Zhengzhou 450002, China)

Abstract: 【Objective】This study proposes an improved YOLOv5s algorithm model for maize trumpet - stage detection. The primary goal is to construct a lightweight model network structure while simultaneously augmenting the detection accuracy, which is crucial for enhancing the efficiency of precise maize spraying and mitigating environmental pollution.【Method】In the context of the YOLOv5s network structure, the FasterNet structure is first utilized to substitute the C3 structure within YOLOv5s, thereby attaining model lightweighting. Subsequently, an in - depth exploration is conducted regarding the influence of Squeeze-and-Excitation (SE) and Convolutional Block Attention Module (CBAM) attention mechanisms on model performance when they are positioned at diverse locations within the backbone network. Finally, a comparative analysis is carried out on the impacts of CIoU, EIoU, and WIoU loss functions in YOLOv5s on model precision.【Result】The experimental results of attention mechanisms comparison reveal that the CBAM and SE attention mechanisms can enhance the model detection precision at specific positions. The highest precision can reach 89.1% and 88.8% respectively. However, when both are incorporated into the backbone network simultaneously, the precision unexpectedly declines to 84.6%. In contrast to CIoU and EIoU, the application of the WIoU loss function is more effective in promoting model performance, with the precision and recall reaching 87.5% and 89.7% respectively. After the improvement, the *mAP*@0.5 of the model on the detection dataset attains 92.3%, which represents a 4.5% increment compared to the original YOLOv5s (87.8%). Additionally, the number of parameters is reduced by 18% and the computational cost is decreased by 21%, while maintaining a relatively rapid detection speed.【Conclusion】The improved algorithm demonstrates excellent performance in

收稿日期：2024-05-22
基金项目：中国博士后科学基金第73批面上资助（2023M730993）；河南农业大学博士启动金（30501158）
作者简介：李文强（1993—），男，河南焦作人，讲师，博士，主要从事智慧农业方向研究，E-mail:272030952@qq.com

the maize trumpet - stage detection task and is capable of providing support for efficient and accurate spraying decision - making under the constraints of limited computing resources. This research not only offers an innovative solution for agricultural precision spraying technology but also lays a solid technical foundation for the development of future intelligent agricultural equipment.
Keywords: Corn Whorl; Object Detection; YOLOv5s Algorithm; Precision Spraying

玉米是全球范围内重要的粮食和饲料作物。2022-2023年，中国玉米播种面积约为4.037×10^{12} m^2[1]，产量达27 720.3万t[2-3]，是中国乃至全球范围内极为重要的粮食来源[4]。然而，传统的玉米喷药方式农药利用率较低，仅不到1%[5]。精准农业技术，特别是基于图像识别的精准喷药技术，可以有效提高农药利用率，减少环境污染，并提高作物产量和品质。近年来，深度学习和卷积神经网络的发展显著提升了机器视觉技术在精准农业中的应用。基于图像识别的精准农业技术通过传感器、控制器或机器视觉对单株作物进行识别和定位，能够实现从无差别喷药到定点喷药的精细化管理[6]。基于图像识别技术的田间植株护理作业不仅可以最大程度地减少环境足迹和降低污染，还能够提高资源利用率和粮食生产效率[7]，为提高玉米生产效率和保护环境提供了一个创新的解决方案。然而，传统图像识别算法在农业边缘设备上的部署面临计算资源的限制，其作业的准确性直接受限于图像识别算法的精度。因此，开发适合农业边缘设备的高效、精准图像识别算法，是实现这一技术潜力的关键。

随着深度学习和卷积神经网络的发展，机器视觉技术的精度得到大幅提升，极大地推动了其在精准农业中的应用，进而革新了传统的农作方式。其中，YOLO（you only look once）作为一种基于卷积神经网络的目标检测算法，凭借其实时性和高效性，成为精准农业中常用的技术之一。吴叶辉等[8]提出一种基于Swin-Tiny Transformer随机增强轻量级模型的图像识别方法对玉米病害图像进行识别，准确率达到了93.5%。郭小燕等[9]提出了一种轻量级YOLOv5s农作物虫害目标检测模型以解决在样本数量不足的情况下农作物虫害识别的问题，对于水稻、玉米平均精度为91.31%。刘冰杰等[10]对YOLOv3、YOLOv5、SSD模型进行对比分析，验证了YOLOv5模型更适合用于作物田间精确除草的自动化作业。然而，基于卷积神经网络上述识别算法在提升检测模型精度的同时，不可避免地导致了模型的复杂并引入了大量的参数，这不利于后期的模型部署和应用。因此，许多研究者着眼于对卷积神经网络进行轻量化处理。何全令[11]提出了一种基于轻量级网络模型的实时目标检测算法GBC-YOLOv5s，模型大小与YOLOv5s相比减少了77.39%，*mAP*提高了2.1%，并且将其在嵌入式除草机器人上进行了实际部署。吴利刚等[12]将以YOLOv5模型为基础，结合轻量级网络ShuffleNetV2，在保证精度的同时，显著降低了模型的运算量和网络层数。王金鹏等[13]针对自然环境下油茶果存在严重遮挡、近景色、小目标等现象提出使用MPDIoU作为YOLOv8n的损失函数，并调整网络向其中加入小目标检测层，最后使用SCConv作为特征提取网络使模型轻量化，为实现自动化采摘提供了必要条件。其次对于精准农业发展，实现精准打药是其关键步骤，张谋等[14]提出了一种可搭载于小型农药喷洒机器人的视觉识别喷药系统，实现对单株玉米秧苗的精确喷药。刘江川[15]提出了一种基于改进CenterNet的玉米雄蕊识别模型。该模型通过分析玉米雄蕊的尺寸分布，简化了骨干特征提取网络，并添加了位置坐标，从而提高了定位精度，降低了雄蕊漏检率，显著提高了玉米授粉的效率。其检测结果分别比Faster R-CNN和YOLOv4模型高出26.22%和3.42%。

基于卷积神经网络的图像识别算法虽然已成功应用于农业，但在算力资源有限的嵌入式平台上，部署复杂且大型的检测模型仍面临困难。因此，本研究提出了一种结合FasterNet网络和注意力机制的改进YOLOv5s算法，用于玉米喇叭口的检测。首先，将YOLOv5s的C3模块替换为轻量级的FasterNet模块，以降低模型复杂度。随后，为了弥补轻量化带来的精度损失，在不同网络层中引入SE和CBAM注意力机制以提高模型性能。最后，使用WIoU损失函数替代CIoU损失函数以增加网络对普通质量锚框的关注度

从而提高检测精度和效率。试验结果表明，改进后的算法在检测精度、参数量和运算量方面均有显著提升。

1 YOLOv5s改进

1.1 YOLOv5s网络模型

YOLOv5[16]是YOLO系列的第5个版本，其在实时目标检测领域具有显著的准确性和效率。YOLOv5s是YOLOv5中最小的模型。其主要包括输入端、主干网络、颈部网络和预测网络四个部分，如图1所示。在输入端，YOLOv5s主要采用了Mosaic数据增强技术，该技术通过将四张不同的训练图片进行随机缩放、裁剪和拼接，从而提高了模型对不同尺度、形状和背景的目标的检测能力。在主干网络部分，YOLOv5s采用了Focus结构和CSP结构。Focus结构通过切片操作将输入图片分割成四个部分，并重新排列成一个新的特征图，这样做可以在不增加计算量的情况下提高特征的丰富性。CSP结构则通过将输入特征图分为两个部分，并分别进行处理后再合并，从而减少了计算量和参数数量，同时保持了特征的丰富性。在颈部网络采用了FPN+PAN结构。FPN结构通过构建不同尺度的特征金字塔，使得模型能够检测到不同大小的目标。PAN结构则进一步优化了特征的融合方式，提高了特征的利用效率。在预测网络部分，YOLOv5s采用了CIoU-Loss作为损失函数。其设计考虑了预测框和真实框之间的面积差异、中心点距离和纵横比差异，能够对目标进行有效识别，这对于检测不同形状和方向的玉米喇叭口尤为重要。

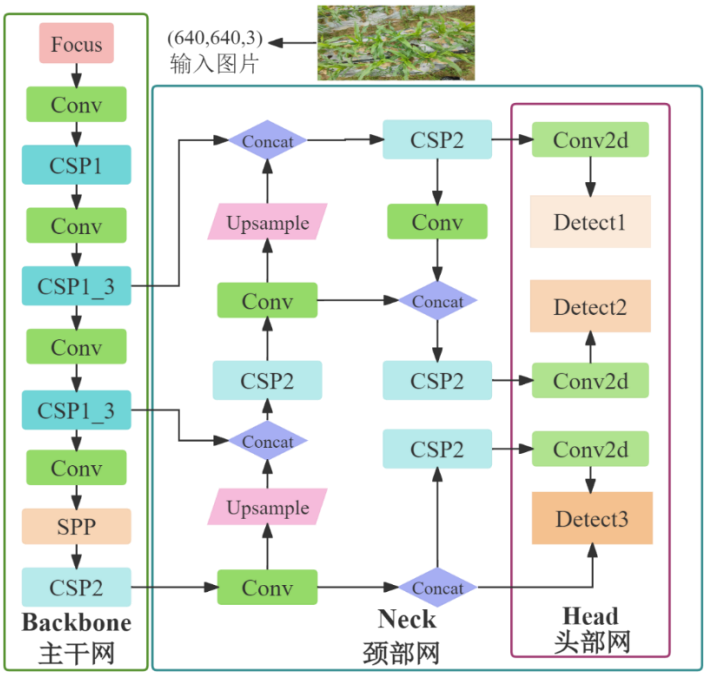

图 1 YOLOv5 网络结构

Fig.1 YOLOv5 Network Structure

YOLOv5s网络模型在目标检测任务上展现了卓越的性能，但其架构中集成了诸如Focus结构、CSP结构、FPN和PAN等复杂组件，导致神经网络在进行多层级、不同尺度的卷积运算过程中产生大量冗余特征图。这种冗余不仅对模型的检测速度产生不利影响，而且由于YOLOv5s的计算图中包含大量操作与节点，进一步加剧了模型的复杂性。因此，对YOLOv5s的特征提取网络模块进行轻量化改造显得尤为重要，以确保在嵌入式设备上部署的玉米喇叭口检测模型能够维持较高的实时性。

1.2 轻量化主干提取网络

在YOLOv5s算法的主干网络中，C3模块的设计有助于提升特征的抽取效率，但同时也带来了显著的内存访问量，进而导致了处理速度的延迟和计算速度的下降。为了满足玉米喇叭口目标检测模型对低延迟和高吞吐量的需求，本研究采用了FasterNet[17]轻量级卷积神经网络架构，其由4个层次阶段和3个主要部分组成，如图2所示，每个阶段都有1组FasterNet模块，并在其前面有1个嵌入或合并层。最后的3个层用于特征分类。在每个FasterNet块内部，包含1个Partial Conv（PConv）层和2个1×1Conv层[18]。

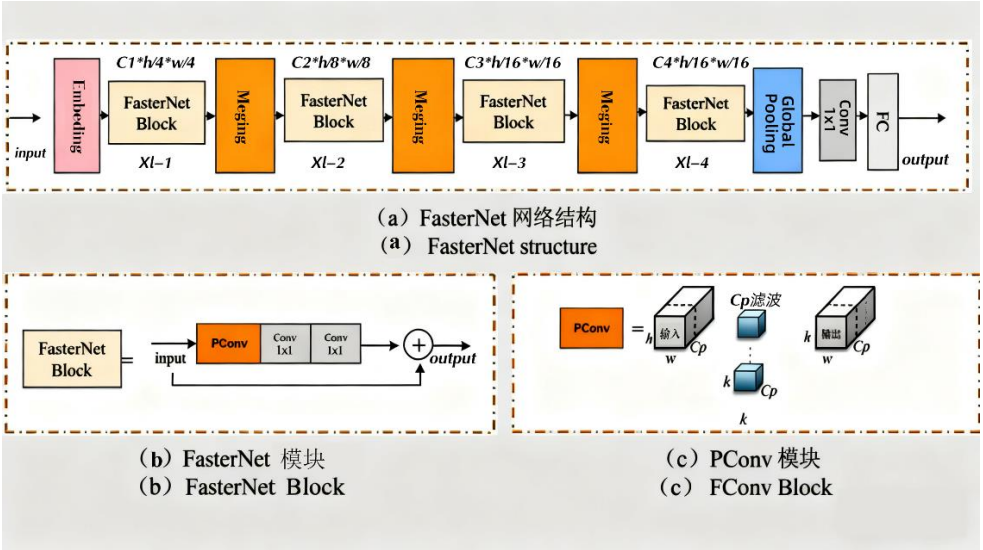

图2 FasterNet整体结构图

Fig.2 FasterNet Overall Structure

这些块共同构成了一个倒置残差块，其中中间层具有扩展的通道数量，并且通过Shortcut连接重用输入特征，并在中间层之后放置标准化和激活层，以保留特征的多样性并实现更低的延迟。PConv层是Fasternet网络中用于处理点云数据的一种卷积层，主要使用Circular Partitioning（CP）点云数据处理滤波技术，其将点云数据划分为多个局部邻域，并在这些邻域上应用卷积滤波器，可以有效提高点云数据的处理速度和效率。这种结构设计使得FasterNet在多种视觉任务中展现出了优异的性能，同时在保持速度的同时，压缩模型大小，提升对玉米喇叭口的特征表达能力，符合嵌入式设备对玉米喇叭口目标检测模型的轻量化的要求。

基于上述FasterNet的主要思想，并结合YOLOv5s中的C3模块，本研究将轻量化的FasterNet主干网络与YOLOv5s算法中的C3模块相融合，采用FasterNet中的局部卷积PConv操作来替代C3模块中原有的卷积层，如图3所示，并命名为C3-Faster网络结构。

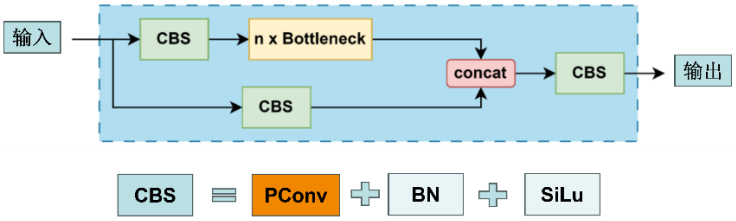

图3 C3-Faster网络结构图

Fig.3 C3-Faster Network Structure

C3-Faster网络架构能够消除冗余的计算和内存访问，实现了对空间特征的高效提取。这种设计策略不仅显著降低了浮点运算（Flops），而且能有效地减少了处理延迟，从而实现了更快的计算速度，这有利于后续的硬件部署。

1.3 注意力机制

1.3.1 SE注意力机制 深度学习中的注意力机制可以帮助人们在有限的资源下，从大量无关背景区域中筛选出具有重要信息的目标区域，帮助人们更高效地处理视觉信息，因此可以借助特定的注意力机制来实现对于玉米喇叭口目标对象的重要信息的处理。常用

的经典注意力机制，有CBAM注意力机制[19]、SE注意力机制[20]、Multi-Scale Dilated Attention（MSDA）多尺度空洞注意力机制[21]等。其中SE注意力机制是一种用于计算机视觉任务中的神经网络架构组件，特别是对于目标检测任务，该机制能够显著提升模型的性能。SE注意力机制通过模型学习到的通道依赖关系来增强特征表示，使得模型能够更加关注信息量大的特征，并抑制不重要的特征。其网络结构如图4所示。

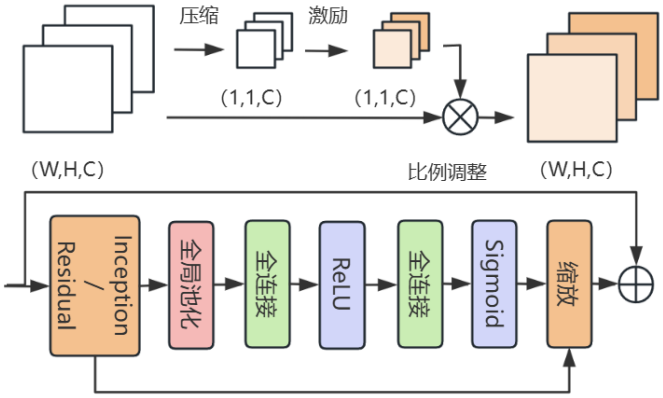

图4 SE模块结构图
Fig.4 SE Module

SE注意力机制旨在通过生成全局响应来增强特征表示。具体来说，对于每个特征图，首先计算一个标量值，该值编码了特征图的全局分布信息，如公式（1）所示。随后，通过两个全连接层（或一个带有ReLU激活函数的瓶颈层）来建模特征图之间的依赖关系。这一过程能够动态地调整特征图的重要性，从而在全局范围内提升特征表示的质量和模型的表现。第1个全连接层通常将通道数减少到一个较小的比例，然后第2个全连接层恢复到原始的通道数，如公式（2）所示。使模型可以学习到不同通道的重要程度。在第2个全连接层之后通常使用Sigmoid激活函数，以生成1组介于0和1之间的权重，这些权重代表了对每个特征图通道的重要性评分。最后将Excitation阶段计算得到的权重应用于原始的特征图上，如公式（3）所示。具体来说，就是将每个通道乘以其对应的权重，这样就能根据重要性对特征进行重新标定。

$$s_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{H} X_{c,i,j} \tag{1}$$

$$e = \sigma(W_2 \cdot \delta(W_1 \cdot s)) \tag{2}$$

$$\tilde{x}_{c,i,j} = e_c \cdot x_{c,i,j} \tag{3}$$

式中：s_c为通道的全局响应值；$X_{c,i,j}$为在通道c上位置(i,j)处的特征值；e为激励阶段计算得到的权重；σ为Sigmoid激活函；δ为表示ReLU激活函数；s为表示通道的全局响应值；$\tilde{x}_{c,i,j}$为调整后的特征值；$x_{c,i,j}$为原始特征值。

因此，对于本研究SE注意力机制可以对玉米喇叭口的喇叭口形状和纹理的通道关键特征进行强化，同时抑制与检测任务无关的背景噪声。集中关注难以检测的区域，提高模型对小型和模糊目标的检测能力，在玉米生长的不同阶段，喇叭口可能会被叶子遮挡或呈现不同的形状变化。SE注意力机制可以使模型更加关注遮挡区域和非刚性变形部分，从而提高在这些复杂场景即不同的光照、天气和土壤条件下呈现不同的外观情况下的检测性能，帮助模型适应这些变化，提高检测的鲁棒性。

1.3.2 CBAM注意力机制 SE注意力机制主要聚焦于通道维度的信息，在某些特定任务中，可能不足以实现精确的特征提取。因此，本研究同时引入了CBAM注意力机制对玉米喇叭口检测模型进行改进，其融合了通道注意力模块（Channel Attention Module, CAM）和空间注意力模块（Spatial Attention Module, SAM），通过这种综合的注意力机制，选择出能够更有效地识别和突出玉米喇叭口的重要特征。

CBAM注意力机制的结构如图5所示，其首先通过通道注意力模块对通道间的依赖性进行建模，为每个通道分配权重，强化了重要的特征表达。随后，空间注意力模块进一步对这些加权后的特征图进行分析，关注特征图中的空间区域，增强了特征的空间表达能力。通过这种层次化的注意力机制，CBAM能够同时考虑通道和空间两个维度的信息，从而更全面地捕捉到图像的关键特征，有效地提高模型在捕捉和识别不同尺寸和形状的玉米喇叭口的性能。

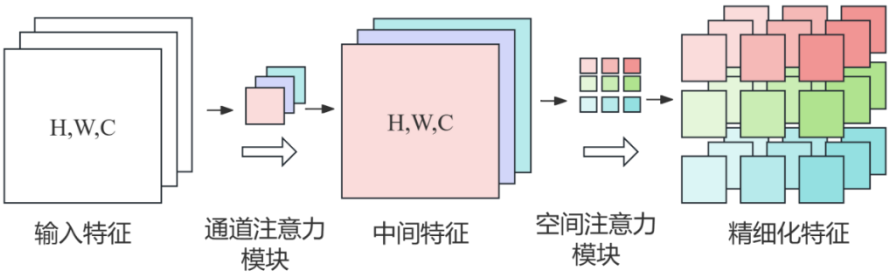

图5 CBAM注意力机制图
Fig.5 CBAM Attention Mechanism

1.4 损失函数

1.4.1 CIoU损失函数 YOLOv5s中采用CIoU[22]损失函数评价模型预测值与真实值之间的差异。其设计综合了交并比和距离度量，不仅考虑了预测框和真实框的交集和并集，还考虑了其长宽比和中心点之间的距离，从而更准确地评估预测框的质量。CIoU的引入可以有效地解决目标检测中的类别不平衡和长宽比不平衡问题，提升了检测的准确性和稳定性。其主要作用是用于衡量预测框和真实框之间的差异，并作为优化目标进行模型训练。具体公式如（4）、（5）、（6）所示。

$$Loss_{CIoU} = 1 - CIoU \tag{4}$$

$$v = \frac{4}{\pi^2}\left(\arctan\frac{W_b}{H_b} - \arctan\frac{W_p}{H_p}\right)^2 \tag{5}$$

$$CIoU = IoU - \frac{d^2}{c^2} - \frac{v^2}{(1 - IoU) + v} \tag{6}$$

式中：H_b为真实框的高度；W_p为预测框的宽度；H_p表示预测框的高度；IoU 为预测框和真实框的交并比；d为预测框和真实框中心点之间的欧氏距离；c 为包含预测框和真实框的最小外接矩形的对角线距离；v 为长宽比的一致性度量。

1.4.2 EIoU损失函数 CIoU损失函数在边界框回归中综合考虑了重叠面积、中心点距离和纵横比。然而，其公式中的v项仅反映纵横比的差异，而未能分别考量宽度和高度与其置信度之间的真实差异。因此，在某些情况下，这可能会阻碍模型对相似性的有效优化。针对这一问题，在CIoU的基础上将纵横比拆开，引入了EIoU损失函数[23]，EIoU的惩罚项是在CIoU的惩罚项基础上将纵横比的影响因子拆开分别计算目标框和锚框的长和宽，该损失函数包含三个部分：重叠损失，中心距离损失，宽高损失，前两部分延续CIoU中的方法，但是宽高损失直接使目标盒与锚框的宽度和高度之差最小，使得收敛速度更快。惩罚项公式如式（7）所示。

$$\begin{aligned} L_{EIoU} &= L_{IoU} + L_{dis} + L_{asp} \\ &= 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(w, w^{gt})}{C_w^2} + \frac{\rho^2(h, h^{gt})}{C_h^2} \end{aligned} \tag{7}$$

式中：L_{dis} 为预测框和真实框中心点之间的距离差异；L_{asp} 为预测框和真实框的长宽比差异；$\rho^2(b, b^{gt})$为预测框和真实框中心点之间的欧氏距离；$\rho^2(w, w^{gt})$为预测框和真实框宽度之间的差异；$\rho^2(h, h^{gt})$为预测框和真实框高度之间的差异。

1.4.3 WIoU损失函数 CIoU损失函数的复杂性源于其对边界框之间的距离、重叠区域以及纵横比等多个几何参数的综合考量，这在提升定位精度的同时，也增加了计算负担。为了缓解这些问题，WIoU损失函数[24]当目标边框与锚框高度重合时，适当减弱距离和纵横比等几何因素的影响，从而增强算法的泛化能力，并克服了CIoU在某些场景下的局限性。同时，融合了注意力机制和梯度增益的概念，这进一步促进了目标检测模型训练过程的优化，提高了模型在处理复杂场景时的鲁棒性和准确性。

WIoU损失函数首先根据距离度量构建了距离注意力，得到了具有两层注意力机制的WIoUv1，如式（8）、（9）所示。

$$\mathcal{L}_{WIoUv1} = \mathcal{R}_{WIoU}\mathcal{L}_{IoU} \tag{8}$$

$$\mathcal{R}_{WIoU} = exp\left(\frac{\left(x - x_{gt}\right)^2 + \left(y - y_{gt}\right)^2}{\left(W_g^2 + H_g^2\right)^*}\right) \tag{9}$$

式中：$\mathcal{R}_{WIoU}$为用于处理预测框与真实框之间的距离；W_g为真实框的宽度；H_g 为真实框的高度。

其次参照*Focal-Loss*[25]设计了一种针对交叉熵的单调聚焦机制WIoUv2，有效降低了简单示例对损失值的贡献。这使得模型能够聚焦于困难示例，获得分类性能的提升。如式（10）、（11）、（12）所示。

$$\mathcal{L}_{WIoUv2} = \mathcal{L}_{\mathrm{IoU}}^{\gamma*}\,\mathcal{L}_{WIoUv1}, \gamma > 0 \tag{10}$$

$$\frac{\partial \mathcal{L}_{WIoUv2}}{\partial \mathcal{L}_{IoU}} = \mathcal{L}_{IoU}^{\gamma^*}\frac{\partial \mathcal{L}_{WIoUv1}}{\partial \mathcal{L}_{IoU}}, \gamma > 0 \tag{11}$$

$$\mathcal{L}_{WIoUv2} = \left(\frac{\mathcal{L}_{IoU}^*}{\overline{\mathcal{L}_{IoU}}}\right)^r \mathcal{L}_{WIoUv1} \tag{12}$$

式中：γ 为大于0的正数，用于调整损失函数的权重；r 为调整系数，用于调整损失函数的权重。

最后利用β构造了一个非单调聚焦系数并将其应用于WIoUv1就得到了具有动态非单调FM的WIoUv3。利用动态非单调FM的明智的梯度增益分配策略，WIoUv3获得了更加优越的性能。如式（13）、（14）所示。

$$\beta = \frac{\mathcal{L}_{IoU}^*}{\bar{\mathcal{L}}_{IoU}} \in [0, +\infty) \tag{13}$$

$$\mathcal{L}_{WIoUv3} = r\mathcal{L}_{WIoUv1}, r = \frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{14}$$

式中：r为调整系数，用于调整损失函数的权重；δ为常数参数，用于调整损失函数的权重；α为常数参数，用于调整损失函数的权重。

因此，WIoU相较于传统的IoU损失函数，通过引入权重系数，有效地解决了目标检测中的类别不平衡问题。WIoU的计算方式基于预测框与真实框的交集面积与加权后的并集面积之比。在此计算过程中，权重系数根据目标的类别和大小进行适当调整，从而使得不同尺寸的目标具有不同的权重，以更好地适应多样化场景下的目标检测任务。

1.5 改进后的整体结构

综上所述，本研究在YOLOv5s算法的基础上进行了多项改进，以实现玉米喇叭口检测模型的轻量化、高效性和准确性。具体改进如下：

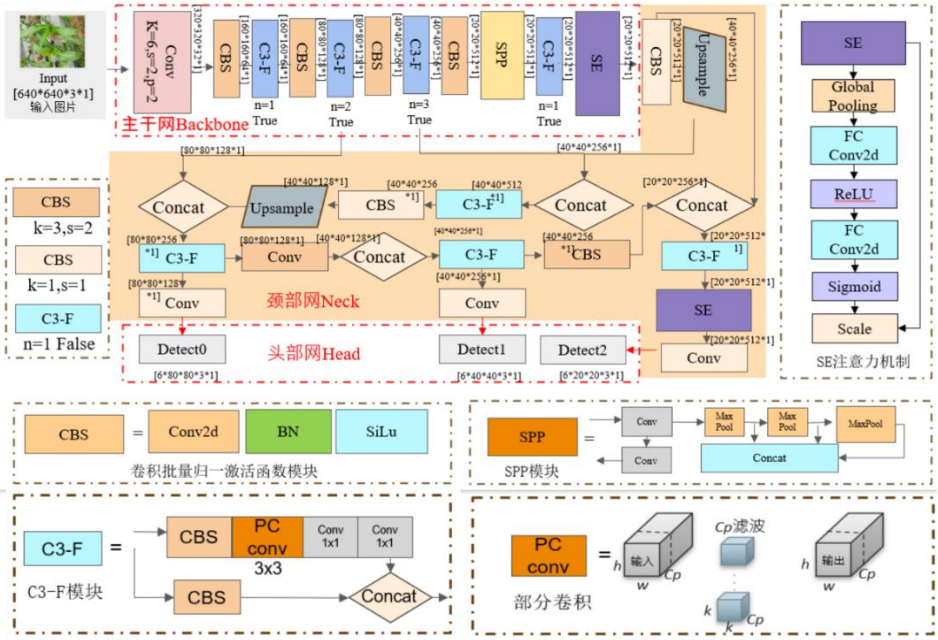

图6 YOLOv5-SE模型架构图

Fig.6 YOLOv5-SE Model Architecture

首先通过将主干网络中的C3模块替换为轻量级的FasterNet基础模块，降低了模型的复杂度和深度。这一替换通过减少计算需求和参数量，提高了模型的简化程度。随后，在Neck部分，引入了C3-Fater模块，该模块由多个卷积层和PConv层组成，并通过拼接（Concat）和上采样（Upsample）操作进一步提取和融合特征。这种紧凑的结构设计有助于提升模型的检测性能。在Head部分，加入了SE注意力机制。SE模块通过全局池化（Global Pooling）计算每个特征图的全局信息，并通过两个全连接层（FC Conv2d）建模特征图之间的依赖关系，最终输出缩放系数来调整特征图的权重。最后，采用WIoU损失函数替代原有的CIoU损失函数，以增加网络对普通质量锚框的关注度，从而进一步提升算法性能。

图6为改进后的YOLOv5模型框图，包括Backbone、Neck和Head三个主要部分。通过这些改进，模型在保持较低计算复杂度的同时，能够实现更高效和准确的目标检测。

2 模型性能试验结果与分析

本研究的数据集由实地场景拍摄获得，捕获的数据集中主要包括生长在中期阶段的玉米。本研究提出的算法在玉米喇叭口数据集上进行了训练和验证。该数据集包含520张图像，其中440张用于训练，剩余80张用于验证。为了进行有效的监督学习，使用Labelimg软件对这520张玉米喇叭口图像进行了精细标注，所有标注信息均遵循YOLO格式，并保存为txt格式文件。

在试验过程中使用GTX2080Ti-GPU作为硬件平台，并采用了PyTorch深度学习框架来执行模型训练。表1详细列出了试验中使用的超参数设置，这些参数经过多次反复调整，以确保模型能够达到最佳的训练效果。

表1 试验超参数配置

Table 1 Experimental Hyperparameter Configuration

参数名称 Parameter Name	取值 Value	参数名称 Parameter Name	取值 Value
学习率 Learning Rate	0.01	迭代次数 Number of Iterations	100

动量 Momentum	0.937	图片大小 Image Size	640*640
权重衰减 Weight Decay	0.000 5	学习率衰减策略 Learning Rate Decay Strategy	cosin
批量大小 Batch Size	16		

2.1 评价指标

为衡量网络结构对玉米喇叭口目标检测的准确性，本研究采用图片检测的精确度（*Precision*，*P*）、召回率（*Recall*，*R*）、均值平均精度（*mean Average Precision*，*mAP*）来评估最终效果。同时为了评估网络在移动端部署的可行性，从以下几个方面对模型性能进行评价：首先，使用单位时间内处理的图像数量（帧率，*FPS*）来衡量模型的检测速度；其次，使用浮点运算数（*FLOPs*）和模型参数量（*parameters*）来评估模型的计算复杂度和规模。其中*mAP*作为整个模型的主要衡量指标，各项指标的具体定义如下列所示。

$$P = \frac{TP}{TP + FP} \tag{15}$$

$$R = \frac{TP}{TP + FN} \tag{16}$$

$$mAP = \frac{1}{n_j \sum_{j=1}^{n_j} AP_j} \tag{17}$$

$$FPS = \frac{1000}{P_{reThime} + InferTime + N_{MSTime}} \tag{18}$$

式中：*TP*（true positive）为预测正确的正样本的数量；*FP*（false positive）为预测错误的正样本的数量；*FN*（false negative）为样本中未被检测出来的正样本的数量；*FPS*是基于batch-size为32的条件下运算的。

2.2 轻量化主干网络结构试验对比分析

将YOLOv5s模型的主干网络中的C3结构替换为C3-Faster结构，经过训练后其性能如表2所示。由表2可知，改进后的轻量化网络YOLOv5s-light的运算量，模型尺寸和参数量分别下降了15%和18%，且模型检测速度有了一定的提升，但准确率降低了2.5%。网络的轻量化在降低模型复杂度的同时，不可避免会导致准确率的下降，因此为了弥补准确率的损失，需要在轻量化网络的基础上进行优化。

表2 网络轻量化前后对比

Table 2 Network Lightweighting Comparison

模型 Model	平均精度均值/% *mAP*	浮点运算次数/G *Flops*	每秒帧数/fps *FPS*	模型大小/Mb *Size*	参数数量/M *Params*
标准模型 YOLOv5s	87.8	15.9	24.4	13.7	7.01
改进主干模型 YOLOv5s-light	85.3	12.6	26.0	11.3	5.78

2.3 注意力机制试验对比分析

如上所述，为了减轻轻量化网络对模型检测精度带来的负面影响，在特征提取网络中引入注意力机制。这一策略的核心思想是通过对玉米喇叭口特征的深入挖掘和强化，提升模型对于这些关键特征的识别能力。然而，注意力机制的应用位置和类型对模型的检测性能有着不同的影响。因此，有必要设计并实施一系列试验，以比较和确定最适合的注意力机制及其最优应用位置。本研究分别将SE注意力机制和CBAM注意力机制添加到特征提取网络的每个C3-Faster模块之后，如图7所示，并分别记录对应的模型为：SE_1/CBAM_1、SE_2/CBAM_2、SE_3/CBAM_3和SE_4/CBAM_4。

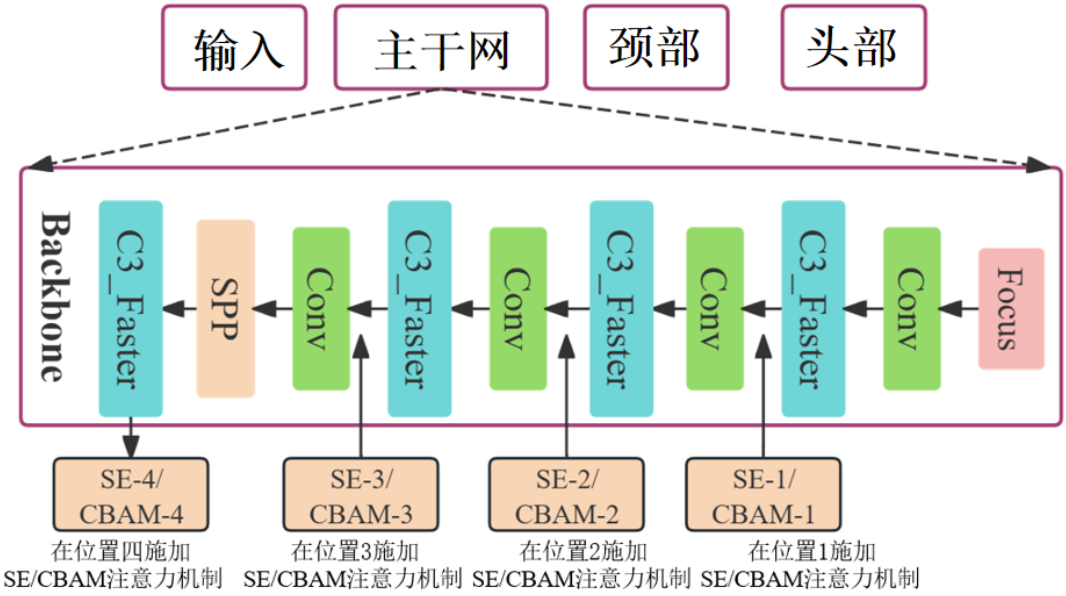

图7 注意力机制添加位置

Fig.7 Attention Mechanism Addition Location

对比不同的注意力机制类型和施加位置对模型检测效果的影响，其试验结果如表3所示。在检测精度方面，将CBAM注意力机制施加在图中四个位置时，模型检测精度分别为85.5%、86.9%、88.9%和89.1%，相较轻量化网络精度分别提高0.2%、1.6%、3.6%和3.8%。与原本YOLOv5s网络精度相比，当CBAM注意力机制添加至位置3和位置4时，模型的精度能够提升；而将SE注意力机制施加在图中四个位置时，模型检测精度分别为87.0%、85.6%、85.6%和88.8%，相较轻量化网络精度分别提高1.7%、0.3%、0.3%和3.5%。与原本YOLOv5s网络精度相比，只有将SE注意力机制添加至位置4时，模型的精度才会提高。这是基于SE注意力模块有选择性地增强有用的特征并抑制不重要的特征的有效机制，从而提高了特征提取的针对性和效果。而CBAM注意力机制由于结合了通道注意力机制和空间注意力机制，更全面地捕捉到重要的特征，提高了特征提取的效率和准确性。

添加CBAM注意力机制的精度提升略高于SE注意力机制。然而，在参数量、运算量和计算速度方面，添加在位置4的SE注意力机制模型表现出最佳性能。此外，在训练过程中，添加CBAM注意力机制的模型表现出较大的波动，这种波动可能源于CBAM对玉米喇叭口数据集的泛化能力不足。综上所述，本研究采用了添加在位置4的SE注意力机制模型。通过这种选择并利用SE注意力机制在特征提取方面的优势，以实现模型轻量化和精度之间的平衡。

试验结果进一步揭示，在轻量化模型的特征提取网络中，较为靠后的位置施加注意力机制能够更有效地提升模型的检测精度。然而，值得注意的是，当SE注意力机制和CBAM注意力机制被同时集成至位置4时，模型的精度意外地下降至84.6%，如表3中CBAM_4+SE_4所示。这一发现说明：注意力的集成并非简单的累积或任意插入，而是

需要根据目标检测网络的具体结构进行细致的选择和定位，以实现注意力机制与模型架构的和谐匹配。

表 3 注意力机制试验对比数据
Table 3 Attention Mechanism Experimental Comparison Data

模型 Model	平均精度均值/% *mAP*	每秒帧数/fps *FPS*	召回率/% *Recall*	参数数量/M *Params*	浮点运算次数/G *Flops*
作用于位置1 CBAM_1	85.5	19.9	84.0	57.8	12.6
作用于位置2 CBAM_2	86.9	19.9	83.7	57.8	12.6
作用于位置3 CBAM_3	88.9	18.8	85.2	57.9	12.6
作用于位置4 CBAM_4	89.1	21.7	88.0	58.1	12.7
作用于位置1 SE_1	87.0	18.0	81.1	57.8	12.6
作用于位置2 SE_2	85.6	17.6	85.2	57.8	12.6
作用于位置3 SE_3	85.6	18.7	82.9	57.8	12.6
作用于位置4 SE_4	88.8	22.2	87.3	57.8	12.6
同时作用于位置4 CBAM_4+SE_4	84.6	22.9	81.7	57.9	12.6

2.4 损失函数试验对比分析

由于玉米喇叭口数据集自身特征并不显著，且存在植株高度差，叶片遮挡，光线不均匀等情况，在轻量化网络结构添加注意力后并不能很好满足实际检测需求，而损失函数的选择对于促进模型的学习优化、提高检测精度和速率至关重要，特别是在模型收敛速度以及小目标检测方面。本研究基于CIoU、WIoU和EIoU的三组损失函数组合，以探讨其对模型训练性能的影响。训练集上的性能比较如表4所示。

根据表4的数据分析，当使用YOLOv5s自带的CIoU损失函数时，模型实现了84.2%的精度和87.3%的召回率，平均精度均值（mAP）达到了88.8%。然而，当改用EIoU损失函数时，虽然精度保持稳定，但召回率有所下降，导致mAP降至88.1%。这一现象可能是由于EIoU损失函数在处理极端长宽比的边界框时，相较于CIoU表现出较低的鲁棒性，以及对尺度变化的敏感性。

表4 损失函数试验对比分析数据
Table 4 Loss Function Experimental Comparison Analysis Data

模型 Model	平均精度均值/% *mAP*	每秒帧数/fps *Flops*	每秒帧数/fps *FPS*	准确率/% *Precision*	召回率/% *Recall*
使用CIoU损失函数 YOLOV5s+Faster+SE+CIoU	88.8	12.6	22.1	84.2	87.3

使用EIoU损失函数 YOLOV5s+Faster+SE+EIoU	88.1	12.6	21.5	84.3	86.9
使用*WIoU*损失函数 YOLOV5s+Faster+SE+WIoU	92.3	12.6	23.6	87.5	89.7

采用WIoU损失函数后，模型的精度提升至87.5%，召回率增至89.7%，*mAP*达到了92.3%。WIoU损失函数的灵活性表现在其能够为不同的IoU区域分配不同的权重，从而提供了一种灵活的调整策略，以适应多样化的训练需求。此外，在图8所示的损失函数对比曲线中，可以观察到WIoU损失函数相较于交并比CIoU和EIoU损失函数，在优化检测框（box_loss）和对象检测（obj_loss）的损失值下降方面表现出显著的优越性，并且更快地趋于稳定。这一结果表明，采用WIoU损失函数的模型在预测结果与实际标注之间的误差较小，模型能够更有效地学习数据中的特征和规律，在区分前景与背景方面表现优异，显著降低了将背景错误分类为对象的概率，并减少了遗漏真实对象的情形，从而有效减少了模型的过拟合现象。

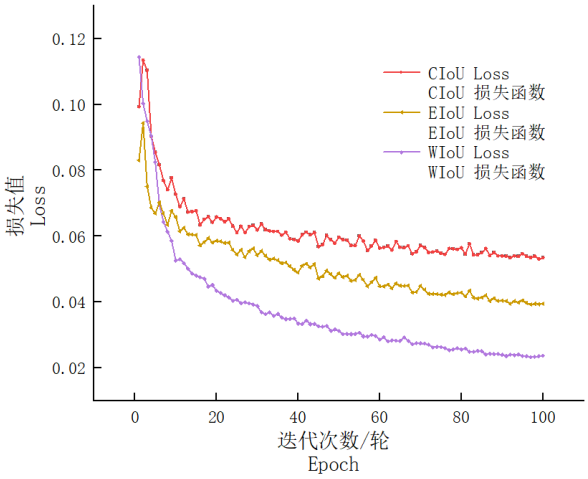

图8 损失函数对比曲线图

Fig.8 Loss Function Comparison Curve

2.5 消融试验对比分析

本研究使用相同的训练、验证数据集分别对改进前、改进后的YOLOv5s算法模型进行训练，同时为了验证各个改进点对模型的影响，本研究进行了消融试验，对比了YOLOv5s、YOLOV5s+Faster、YOLOV5s+Faster+SE和YOLOV5s+Faster+SE+WIoU四种模型，具体试验如表5所示：

通过消融试验的分析，观察到在YOLOv5s基础模型中引入FasterNet轻量化网络结构后，每秒帧数（*FPS*）有所增加，运算量从15.9 *GFlops*降低至12.6 *GFlops*，减少约21%。这一改进归因于FasterNet轻量化网络结构所采用的端到端训练模式和共享特征映射策略，有效减少了计算冗余，提升了模型效率。然而，平均精度均值（*mAP*）下降至85.3%，这主要是由于小目标检测性能的降低所致。

表5 消融试验对比分析数据

Table 5 Ablation Experiment Comparison Analysis Data

模型 Model	平均精度均值/% *mAP*	浮点运算次数/G *Flops*	每10秒帧数/10 fps *FPS*	参数数量/M *Params*	模型大小/MB *Size*

标准模型 YOLOv5s	87.8	15.9	24.4	7.01	13.7
进一步改进Faster主干 YOLOV5s+Faster	85.3	12.6	26.0	5.78	11.3
进一步改进SE注意力 YOLOV5s+Faster+SE	88.8	12.6	22.1	5.78	11.3
进一步改进WIoU损失函数 YOLOV5s+Faster+SE+WIoU	92.3	12.6	23.6	5.78	11.3

为了在保持模型轻量化的同时提高检测准确度，进一步集成了SE注意力机制和WIoU损失函数。SE注意力机制展现了对细粒度特征的显著敏感性，而WIoU损失函数的灵活性和多样性有助于将*mAP*值提升至92.3%，较原始YOLOv5s算法提高了4.5%。因此，本研究提出的每项改进均对模型的轻量化和*mAP*的提升作出了重要贡献，验证了各改进点的有效性。

研究结果表明，本研究所提出的改进算法增强了模型的特征提取能力和特征融合能力，能够更有效地提取玉米喇叭口的特征信息，从而实现了对玉米喇叭口特征的准确检测。这些改进不仅提升了模型的性能，也为农业精准打药技术的发展提供了新的技术路径。图10分别展示了在简单背景和复杂背景下的检测效果。

（A）简单背景
(A) Simple Background

（B）复杂背景
(B) Complex Background

图10 检测结果图
Fig.10 Detection Results

2.6 与其他研究讨论对比及展望

刘冰杰等[10]分析了YOLOv3、YOLOv5和SSD模型，发现YOLOv5更适合作物田间精确除草，但未深入优化其复杂性。郭小燕等[9]对YOLOv5s进行了改进，检测病虫害时实现了2.5个百分点的*mAP*提升，模型参数量为4.95 M。相比之下，改进后的模型参数量为5.78 M，但实现了更高的4.5%提升。吴利刚等[12]基于YOLOv5结合ShuffleNetV2，显著降低了运算量和网络层数，确保了精度。然而，该研究仅使用了SE注意力机制，而本研究引入了CBAM，并探讨了两者的联合使用，从而进一步提升了注意力机制的有效性。在*mAP*提升方面，本研究优于吴的研究，尽管在轻量化方面仍存在改进空间。

综上所述，本研究提出的改进模型在精度和轻量化方面均表现出显著优势，尽管仍有进一步优化的可能性。这些结果表明，本方法在当前研究基础上具有更大的应用潜力和发展前景。

3 结果与讨论

本研究提出了一种基于YOLOv5s模型的改进算法，旨在实现玉米喇叭口的高精度与轻量化检测。该算法在评价指标*mAP*0.5上达到了92.3%的卓越精度，同时将模型体积压缩至11.3 MB，显著超越了传统方法在识别准确率和模型轻量化方面的表现。为了设计出对计算资源匮乏的田间边缘农业设备友好且高效的识别算法，本研究设计并实施了一系列对比试验。首先，通过将传统的C3网络结构替换为轻量级的FasterNet网络结构，该模型实现了高效轻量化。这一策略不仅大幅减少了模型的参数量和运算量，还维持了较高的检测速度，从而为模型在资源受限的移动设备上的部署提供了可行性，极大地拓宽了其在农业实际应用中的适用范围。其次，引入SE注意力机制进一步增强了模型对玉米喇叭口特征的识别能力。SE机制通过强化关键特征通道的权重，有效抑制了背景噪声，提升了模型在复杂环境下的鲁棒性。试验结果显示，SE机制在主干网络的末尾层展现出最佳的性能提升，显著加强了主干的特征提取能力。在损失函数的选择上，采用WIoU损失函数为模型带来了显著的性能提升。与传统的CIoU和EIoU损失函数相比，WIoU在处理小目标和复杂背景下的玉米喇叭口检测时，表现出更高的精度和更快的收敛速度。这一发现不仅验证了WIoU在目标检测任务中的优越性，也为未来类似任务中的损失函数选择提供了宝贵的参考。

综上所述，本研究提出的改进算法在保持模型轻量化的同时，实现了对玉米喇叭口的高精度检测。这一成果不仅为农业精准打药技术的发展开辟了新的技术路径，也为构建资源节约型和环境友好型的农业体系奠定了坚实的科学基础。与刘冰杰[10]的研究相比，本研究在实现高精度检测的同时，进一步优化了模型的轻量化特性。此外，何全令[11]提出的GBC-YOLOv5s虽然在模型大小上取得了显著减少，但在实际应用中可能面临对小目标的检测精度不足的问题。此外，张谋等[14]的研究虽然实现了对整株玉米的识别进而喷洒农药，但未能针对局部区域进行精细化检测。本研究的创新之处在于专注识别玉米易受病虫害侵害的喇叭口局部区域，为精准农业提供了更为细致的视觉支持。尽管在极端遮挡和复杂光照条件下的检测仍面临挑战，未来的研究将致力于进一步优化模型结构，以提升其在这些复杂场景下的适应性和鲁棒性。

参考文献 References:

[1] 齐驰名. 2022年下半年国内外玉米供需形势和行情展望[J]. 中国国情国力, 2022(9): 77-79.
QI C M. Supply and demand situation and market prospect of corn at home and abroad in the second half of 2022[J]. China National Conditions and Strength, 2022(9): 77-79.

[2] 王凯. 玉米三个重要生长期植保方案研究[D]. 泰安: 山东农业大学, 2017.
WANG K. Research on plant protection plans for three important growth stages of corn[D]. Taian: Shandong Agricultural University, 2017.

[3] World of Corn 2023. (2023). United States Department of Agriculture.National corn growers association.[Online].Available:https://www.worldofcorn.com/#world-corn-production.

[4] 国家统计局[EB/OL]. [2021-3-15].https://data.stats.gov.cn/.

[5] 姜朵朵. 典型农药在我国三种粮食产地残留特征及膳食风险评估[D]. 北京: 中国农业科学院, 2021.
JIANG D D. Residual characteristics and dietary risk assessments of typical pesticides in main areas of three types of grains production in China[D]. Beijing: Chinese Academy of Agricultural Sciences, 2021.

[6] DARBYSHIRE M, SALAZAR-GOMEZ A, GAO J F, et al. Towards practical object detection for weed spraying in precision agriculture[J]. Frontiers in Plant Science, 2023, 14: 1183277.

[7] FOUNTAS S, MYLONAS N, MALOUNAS I, et al. Agricultural robotics for field operations[J]. Sensors, 2020, 20(9): 2672.

[8] 吴叶辉, 李汝嘉, 季荣彪, 等. 基于随机增强Swin-Tiny Transformer的玉米病害识别及应用[J]. 吉林大学学报(理学版), 2024, 62(2): 381-390.

WU Y H, LI R J, JI R B, et al. Maize disease recognition and application based on random augmentation swin-tiny transformer[J]. Journal of Jilin University (Science Edition), 2024, 62(2): 381-390.

[9] 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 X Y, YU S Q. A lightweight YOLOv5S crop pest target detection model[J]. Journal of Nanjing Agricultural University, 2024, 47(5): 1009-1018.

[10] 刘冰杰, 周雅楠, 周小辉, 等. 基于深度学习的玉米田间杂草识别模型研究[J]. 河南农业大学学报, 2024, 58(2): 279-286.

LIU B J, ZHOU Y N, ZHOU X H, et al. Deep learning-based weed recognition model in the maize field[J]. Journal of Henan Agricultural University, 2024, 58(2): 279-286.

[11] 何全令. 基于轻量化网络模型的除草机器人目标识别方法研究[D]. 合肥: 安徽农业大学, 2023.

HE Q L. Research on Target Recognition Method of Weeding Robot Based on Lightweight Network Model[D]. Hefei: Anhui Agricultural University, 2023.

[12] 吴利刚, 吕媛媛, 周倩, 等. 基于改进YOLOv5的轻量级黄花成熟检测方法[J]. 中国农机化学报, 2024, 45(7): 235-242, 268.

WU L G, LÜ Y Y, ZHOU Q, et al. Lightweight method for maturity detection of *Hemerocallis citrina* Baroni based on improved YOLOv5[J]. Journal of Chinese Agricultural Mechanization, 2024, 45(7): 235-242, 268.

[13] 刘江川. 基于深度学习的玉米雄蕊遥感监测方法研究[D]. 杨凌: 西北农林科技大学, 2022.

LIU J C. Research on remote sensing monitoring method of maize stamens based on deep learning[D]. Yangling: Northwest A & F University, 2022.

[14] 张谋, 蒲烨锋, 江雯, 等. 基于双目视觉的农药喷洒机器人识别系统[J]. 农业工程, 2023, 13(10): 27-31.

ZHANG M, PU Y F, JIANG W, et al. Identification system of pesticide spraying robot based on binocular vision[J]. Agricultural Engineering, 2023, 13(10): 27-31.

[15] 王金鹏, 何萌, 甄乾广, 等. 基于COF-YOLOv 8n的油茶果静、动态检测计数[J/OL]. 农业机械学报, 2024: 1-15. (2024-01-17). https://kns.cnki.net/KCMS/detail/detail.aspx?filename=NYJX20240115009&dbname=CJFD&dbcode=CJFQ.

WANG J P, HE M, ZHEN Q G, et al. *Camellia* oleifera fruit static and dynamic detection counting based on improved COF-YOLO v8n[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2024: 1-15. (2024-01-17). https://kns.cnki.net/KCMS/detail/detail.aspx?filename=NYJX20240115009&dbname=CJFD&dbcode=CJFQ.

[16] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016. Las Vegas, NV, USA. IEEE, 2016: 779-788.

[17] CHEN J R, KAO S H, HE H, et al. Run, don’t walk: Chasing higher FLOPS for faster neural networks[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 17-24, 2023. Vancouver, BC, Canada. IEEE, 2023: 12021-12031.

[18] MA X L, GUO F M, NIU W, et al. PCONV: The missing but desirable sparsity in DNN weight pruning for real-time execution on mobile devices[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(4): 5117-5124.

[19] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block attention module[C]//Computer Vision – ECCV 2018. Cham: Springer, 2018: 3-19.

[20] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018. Salt Lake City, UT. IEEE, 2018: 7132-7141.

[21] JIAO J Y, TANG Y-M, LIN K-Y, et al. DilateFormer: Multi-scale dilated transformer for visual recognition[J]. IEEE Transactions on Multimedia, 2023, 25: 8906-8919.

[22] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.

[23] LI X, WANG W, HU R, et al. EIOU Loss for Fully Convolutional Instance Segmentation [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 11365-11374.

[24] CHO Y J. Weighted Intersection over Union (wIoU) for evaluating image segmentation[J]. Pattern Recognition Letters, 2024, 185: 101-107.

[25] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.