

基于特征增强和群组混合注意力的棉花病害检测

王宏强 赵 晖 贾振红

新疆大学计算机科学与技术学院 乌鲁木齐 830049

(zsebzy@stu.xju.edu.cn)

摘 要 为了在田间真实环境中对棉花病害进行快速、准确检测,提出了一个基于特征增强和注意力机制的棉花病害目标检测模型。为确保模型在真实田间环境中检测的准确性,在 Neck 模块中,使用 Group Mix Attention 来联系上下文信息丰富特征图信息,捕获更多样化和细微的特征。使用特征增强模块对特征图进行加权处理,以降低背景或其他物体对图像中目标的干扰。所提模型使用了 SIoU 损失函数,能够有效提高真实田间环境中模型的检测精度,有效减少了模型误检和漏检的情况。实验结果表明,所提模型在自建真实田间环境棉花病害目标检测数据集上均表现出色,有效提高了检测精度。与基线模型相比,所提模型的 *mAP* 和 *Precision* 分别提升了 2 个百分点和 4.5 个百分点。

关键词: 棉花病害;小目标;特征增强;注意力机制;损失函数

中图分类号 TP391

Cotton Disease Detection Based on Feature Enhancement and Group Mix Attention

WANG Hongqiang, ZHAO Hui and JIA Zhenhong

College of Computer Science and Technology, Xinjiang University, Urumqi 830049, China

Abstract In order to achieve rapid and accurate detection of cotton diseases in real field environments, this paper proposes a cotton disease target detection model based on feature enhancement and attention mechanism. To ensure the accuracy of the model's detection in real field environments, an improved feature enhancement module is used in the Neck module to weight the feature maps and reduce the interference of background or other objects on the targets in the image. After the feature enhancement module, Group Mix Attention is used to connect contextual information and enrich the feature map information. The proposed model can effectively improve the detection accuracy of models in real field environments, effectively reducing the occurrence of model false positives and false negatives using SIoU loss function. The experimental results show that the proposed model performs well on the self built real field environment cotton disease target detection dataset, effectively improving the detection accuracy of the model in real field environments. Compared with the baseline model, the *mAP* and *Precision* have increased by 2 percentage points and 4.5 percentage points.

Keywords Cotton diseases and pests, Small goals, Feature enhancement, Attention mechanism, Loss function

1 引言

近年来,计算机视觉在农业领域中的应用逐渐兴起,为作物检测提供了新的解决方案。目标检测是计算机视觉领域的一项关键技术,用于对图像中的目标进行分类和定位。目标检测技术主要分为两种类型的检测器:一阶段目标检测器和两阶段目标检测器。一方面,YOLO^[1],SSD^[2],RetinaNet^[3],EfficientDet^[4],FCOS^[5]和 CornerNet^[6]等一阶段目标检测器以其极快的检测速度而闻名,成为实时应用的理想选择。另一方面,两阶段目标检测器,如 R-CNN^[7],Cascade R-CNN^[8],Fast R-CNN^[9]和 Faster R-CNN^[10],在实现高检测精度方面表现出色,但可能相对较慢。

这些技术不仅能够大幅提升检测速度和精度,还能实现自动化和智能化检测,减少对人力的依赖。基于深度学习的

目标检测模型^[11-14]不仅能够找到图像中叶片病斑的位置,还可以对病斑进行分类,处理多种病害。如果能够精确地检测作物上的病害,将有助于制定农药剂量,精准施用农药,并提高作物的质量和产量。因此,越来越多的研究人员开始将目标检测技术应用于作物病虫害检测领域。起初,Syed-Ab-Rahman等^[15]使用 Faster R-CNN 目标检测模型完成了对简单背景下病虫害的检测。然而,在真实田间环境中,采集到的图像可能存在背景复杂、叶片密集互相遮挡以及复杂的天气条件等因素,这些问题可能会影响模型的检测性能。该模型在实际复杂背景下的表现有待进一步验证和优化。

为了应对这些挑战,研究人员开始针对真实田间环境下的病害目标检测进行研究,以自然环境下拍摄的病害叶片图像为数据集^[12,16],对检测模型进行训练,从而提高目标检测模型在复杂环境中的性能。真实田间环境下的图像包含大量

基金项目:自治区重点研发项目(2023B01032);国家重大科技专项课题(2022ZD0115802)

This work was supported by the Key R&D Projects in the Autonomous Region(2023B01032) and National Science and Technology Major Project of the Ministry of Science and Technology of China(2022ZD0115802).

通信作者:赵晖(zhaohui@xju.edu.cn)

作物和其他物体。在这些图像中,识别和定位极小且密集的危害斑点目标为检测器带来了巨大的困难。

一些研究^[17-19]尝试采用多尺度特征融合的方法增强特征表示,融合低分辨率和高分辨率特征级别,以丰富其细节和语义信息,根据对象的不同尺度大小在不同层次上检测物体,减少了因叶片密集分布和小目标害虫导致的误检和漏检问题。但是当背景中有与病斑相似的纹理时,边界框的回归精度不高,往往无法达到良好的检测性能。另外,许多研究人员^[20-21]尝试引入各种注意力机制实现关键特征提取,通过自适应调节每个尺度特征的权重,使网络在获得不同尺寸特征时更好地将其融合,以进一步强化对象特征,提高模型的前景和背景辨别能力,减少了小型病斑和相似纹理的病斑容易误检和漏检的问题。然而,由于注意力机制过度关注感兴趣的区域,所以部分区域信息被忽视,导致某些小目标对象被漏检。

最近,对专用目标检测任务的需求不断增长,例如密集目标检测和小型目标检测。虽然在应对这些挑战方面已经取得了一些进展,但已有研究仍然存在重大差异。本文探讨了这些持续存在的挑战,分析了棉花叶片病害在真实田间环境下的图像中的影响。在真实田间环境中拍摄的棉花图像中,棉花叶片密集且互相遮挡,某些病害的病斑与背景信息或其他物体纹理相似,同一病害发展程度不同也会导致病斑尺度差异较大。基于这些问题,我们提出了一种使用 YOLO 框架的棉花病害检测模型,该模型侧重于检测棉田中的 4 种常见病害(轮纹斑病、黑斑病、叶霉病和白线病)。我们通过特征增强和注意力对模型进行了技术改进,使模型能够为各种疾病获得更独特的特征。这些改进使该模型更适合在复杂环境中检测疾病。本文的贡献如下:

1) 构建一个具有复杂田间背景的棉花病害图像数据集,其涵盖 4 种类型的棉花病害,包括多叶片、多角度和多种光照条件下的混合拍摄方法。

2) 在模型中添加 Group Mix Attention^[22]和特征增强模块(FEM),并更换损失函数,以改进 YOLO 模型,减少误检。

3) 与基线模型相比,我们提出的方法在自建数据集上实现了 90.4% 的精确度和 86.6% 的 mAP,性能显著提高,优于其他先进的目标检测模型。

2 相关工作

近年来,针对复杂田间环境下目标检测的研究主要围绕特征重构、多尺度特征融合和注意力机制三大技术方向展开探索。基于特征重构的策略致力于提升目标与背景的可分能力,典型工作如 Deng 等^[23]设计了一个名为特征纹理传输(FTT)的新模块,用于超解析特征并同时提取可靠的区域细节,使用交叉分辨率蒸馏机制传递感知细节能力,并设计前景-背景-平衡损失函数以缓解前景和背景面积不平衡的问题。Li 等^[24]通过引入上采样层重构特征融合网络,增加对小目标特征的关注,并引入 SPD 卷积构建模块,提升模型特征提取能力,利用 EIOU 损失函数减少训练过程中的位置损失。这类方法虽能改善纹理混淆现象,但对前景与背景的语义关联建模不足,在动态场景下的泛化能力仍有局限。

基于多尺度特征集成方法聚焦于解决目标尺度差异问题。Jing 等^[25]提出了一种在低分辨率下检测小物体的算法,

通过特征融合模块丰富对象的语义信息和特征表示,并利用视觉感知模块捕获感兴趣的实例,特征编码模块扩展了感受野,提高了模型的语义理解能力,使其更加适应多变环境。

为了解决图像中目标尺度差异大和目标密集且较小的问题,Gao 等^[26]设计了一个多尺度特征金字塔,实现了各级特征信息的高效集成。其使用残差特征增强模块,来增强对不变比例上下文信息的提取,减少金字塔网络中特征图最高层的信息丢失,为小目标检测提供了更丰富的颗粒特征。该方法为了缩放特征图,有效利用了同一空间位置不同通道的特征信息。此类方法普遍面临特征层间对齐效率低、浅层细节与深层语义融合不充分等问题,且在特征传递过程中细粒度信息易丢失。

注意力驱动的方法侧重于提升特征选择能力。Ma 等^[27]提出的 YOLOv5-lotus 方法,利用 CA 注意力机制捕获空间选择性注意力图,加强模型信道间的关系,使得不同信道的特征能够更好地交互和融合,抑制不相关或干扰信息,精确捕捉小目标的位置信息,实现高检测精度和精确定位预测。Song 等^[28]提出了一种基于多尺度混合注意力的检测器,通过多尺度混合注意力在多尺度要素上丰富大型对象和小目标对象之间的联系,增强了不同层次的语义信息,减少了背景信息的干扰。Huang 等^[29]通过引入 ECA 通道注意力机制增强目标的散射特征,抑制不相关的背景信息,并通过滑动窗口分区方法获得全局感知能力,减少小目标漏检问题。现有注意力机制仍存在显著缺陷:全局注意力易受复杂背景干扰,缺乏对微小目标的定向聚焦能力,且动态权重分配与目标尺度适应性不足。

现有方法尽管在基准测试中取得进展,但在真实复杂场景中仍面临本质性挑战:1) 当目标与背景纹理高度相似时,特征判别机制易失效;2) 多尺度特征融合过程中,小目标的关键细节持续衰减;3) 注意力机制对微小目标的关注度与大目标存在显著差距,且计算效率亟待优化。本文针对这些问题,提出融合群组混合注意力(GMA)与特征增强模块(FEM)的改进模型,旨在提升复杂环境下棉花病害检测的鲁棒性。

3 本文方法

3.1 整体概述

YOLO^[30]是一种基于回归的单级对象检测算法,由 3 个主要部分组成:骨干、颈部和头部。最初,图像被输入到网络中。主干通过 Focus, Conv, C3 和空间金字塔池(SPPF)从图像中提取高级特征。然后,这些特征穿过颈部,丰富了特征图的表现力,增强了对大小差异很大的目标的适应性。Neck 通过特征金字塔网络(FPN)和金字塔注意力网络(PAN)整合多级特征数据,以提高特征表达能力,丰富特征信息;然后将得到的图像特征传输到头部。Head 通过 NMS 过滤边界框,并输出边界框及其相应的类别作为最终的对象检测结果。这种多尺度检测策略使 YOLO 能够处理各种大小的复杂场景和目标,稳定且准确地完成目标检测任务。

在棉花病害数据集的收集过程中,发现了棉花病害检测中的许多问题,如目标对象与背景纹理的相似性、不规则的病斑区域以及病害点之间的相似性。我们建议使用 YOLO-GFM 在复杂背景下检测棉花病害,以解决这些问题,如图 1 所示。我们在颈部的 Neck 层中将 C3 模块与 Group Mix At-

tention 模块融合,然后加入了特征增强模块,并引入了 SIoU 损失函数,使模型能够捕获更多关键特征。

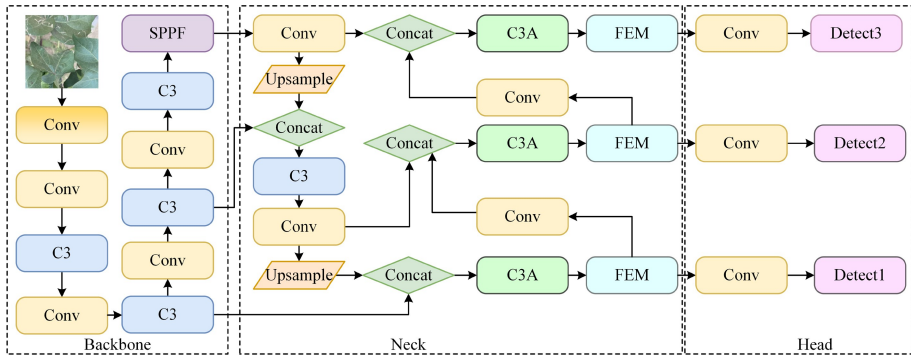


图1 YOLO-GFM 模型

Fig.1 YOLO-GFM model

3.2 Group Mix Attention

在棉花病害检测中,不同病害的病斑存在相似表现,会导致模型检测性能降低。我们引入了 Group Mix Attention 模块,通过获得的上下文信息建立准确的全局关系,捕获更多多样化和丰富的特征,使模型能够学习更全面的对象特征并提高其判别能力。

Group Mix Attention 模型结构如图 2 所示。

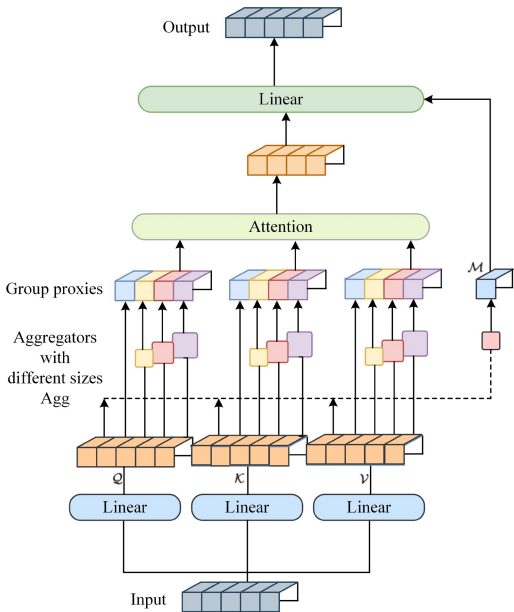


图2 Group Mix Attention 模型

Fig.2 Group Mix Attention model

首先,使用 DepthWiseConv2D 对输入特征进行处理,然后与原始输入特征进行残差计算,以获得具有位置权重的特征图。然后,它们通过线性层映射到查询(Q)、关键字(K)和价值(V)向量。Q, K和V被平均分为 5 个部分,一个部分使用恒等映射保留为M,而其他 4 个部分被发送到聚合器进行注意力计算。聚合器由多个可分离的卷积层、归一化层和激活层组成,每个分支都能够独立处理不同的特征,并在不同的尺度上捕获信息。

接下来对K向量进行 softmax 运算,然后与V向量进行点积,以获得每个关键向量的注意力权重。这些权重表示Q和K之间的相关性。通过点积计算,结合Q向量和注意力权重,生成高效的注意力G,以确保注意力机制能够有效地捕捉和利用输入特征之间的相关性,提高注意力计算的效率。

此步骤按式(1)执行:

$$G = Q \otimes \text{Softmax}(K^T) \times V \quad (1)$$

同时,对V向量执行 DepthWiseConv2D 计算,并与Q向量进行点积运算,获得卷积相对位置编码(ConvRE)。卷积相对位置编码可以捕获输入特征的空间位置关系,使模型能够更好地理解和利用输入特征的空间结构。

$$\text{ConvRE} = Q \odot \text{DepthWiseConv2D}(V) \quad (2)$$

在合并和整形步骤中,对高效注意力G和卷积相对位置编码ConvRE的结果进行加权、合并,并与M连接,以平衡不同特征的贡献,从而提高特征表示的整体质量,提高特征表示的多样性和完整性。最后,通过 Linear 和 Dropout 对最终的注意力特征进行处理,以提高输出特征的稳定性和鲁棒性。

$$\text{ConvAttn}(X) = \text{concat}(G + \text{ConvRE}, M) \quad (3)$$

3.3 特征增强模块

在棉花病害检测中,某些病害的病斑与背景或其他对象纹理相似,使得模型无法完全捕捉到这些对象的前景和背景判别信息。我们使用了特征增强模块来学习更多的判别特征,并提高特征图的表达能力。特征图分为 3 个主要分支,如图 3 所示。

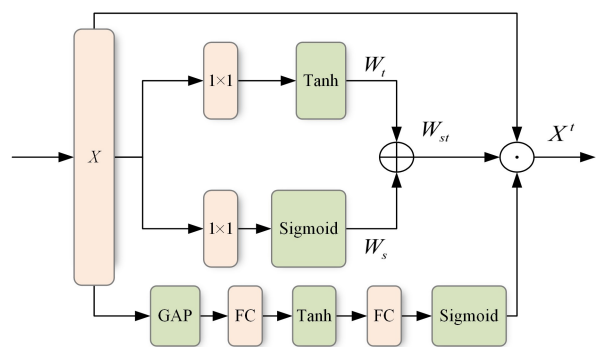


图3 特征增强模块

Fig.3 Feature enhancement module

在第一个分支中,计算得到权重图Wt, Wc能够在通道层面选择的关键特征。

$$W_c = FC(FC(GAP(X)))$$

其中,GAP(·)表示全局平均池化操作;FC(·)是全连接层——前一个全连接层使用 ReLU 激活函数,后一个全连接层使用 Sigmoid 激活函数。

在第二个分支中,将输入特征图再分为两个子分支,使用两个非线性激活算子 Tanh 和 Sigmoid 函数,帮助模型有效地

学习和区分物体与其周围背景或其他物体。输入的特征图 $X \in R^{H \times W \times C}$ 经过一个 1×1 卷积后, 一个分支通过 Tanh 激活函数, 另一个分支通过 Sigmoid 激活函数, 经过计算后分别得到权重图 $W_t \in [-1, 1]^{H \times W \times C}$ 和 $W_s \in [0, 1]^{H \times W \times C}$ 。其中, W_t 用于增强病斑和背景纹理之间的特征差异, 而 W_s 用于学习有关病斑的更多语义信息。这里, X 是头部的特征图。软权重映射 W_s 和 W_t 通过投影变换 F 预测得到, 如式(4)所示:

$$F: X \rightarrow W_i, i \in t, s \quad (4)$$

注意, 当特征图 X 中相应位置的权重图 W_s 的元素值趋于 1 时, 表示具有更多的对象信息。元素值 W_t 趋向 1 表示前景存在, -1 表示背景存在。

然后, 融合 W_s 和 W_t , 通过元素求和得到新的权重图 $W_{st} \in [-1, 2]^{H \times W \times C}$ 。融合权重图 W_{st} 可以引导模型学习更多的语义信息。

$$W_{st} = W_s + W_t \quad (5)$$

在第三个分支中, 在权重图 W_{st} , W_t 和输入特征 X 上进行元素点积操作, 以获得最终的输出特征图 X' 。

$$X' = W_{st} \odot W_t \odot X \quad (6)$$

特征增强模块可以通过两个非线性激活算子 (Sigmoid 和 Tanh) 有效地区分前景和背景信息。融合权重图 W_{st} 可以显式保留对象信息, 同时隐式包含背景信息。 W_t 可以选择渠道层面的主导特征, 以及 W_s 和 W_t 的协作, 使得 X' 变得更具辨别力。特征增强模块增强了模型的学习能力, 使其能够学习更多的判别特征。

3.4 损失函数

在棉花病害检测中, 同一病害在不同发展阶段的病斑形状与轮廓差异较大。为了确保检测的可靠性, 使用 SIoU 损失^[31]代替 CIoU 进行模型的反向传播。SIoU 损失考虑了目标框的位置、大小和形状信息, 不仅依赖于边界框回归指标的聚合, 还引入了向量角度来解决传统方法中预测框漂移导致的收敛速度慢、效率低的问题, 使模型能够更好地定位目标, 提高了算法的准确性。

SIoU 损失包括 4 个成本函数: 角度损失、距离损失、形状损失和 IoU 损失。首先, 添加角度损失, 以快速引导预测框沿最近的 x 轴或 y 轴朝向目标框移动; 然后, 逐渐接近真实框。

$$\begin{cases} x = \frac{C_h}{\sigma} = \sin(\alpha) \\ \sigma = \sqrt{(b_{cx}^{gt} - b_{cx})^2 + (b_{cy}^{gt} - b_{cy})^2} \\ \Delta = 1 - 2 \times \sin^2\left(\arcsin(x) - \frac{\pi}{4}\right) \end{cases} \quad (7)$$

其中, $C_h = \max(b_{cy}^{gt}, b_{cy}) - \min(b_{cy}^{gt}, b_{cy})$, 如图 4(a) 所示。在训练过程中, 如果 $\alpha \leq \frac{\pi}{4}$, 则 α 最小化, 反之 β 最小化。当中心点在 x 轴或 y 轴上对齐时, $\Delta = 0$ 。当中心点连接到 x 轴 45° 时, $\Delta = 1$ 。考虑到新定义的角度成本, 距离损失重新定义为:

$$\begin{cases} \gamma = 2 - \Delta \\ \Delta = \sum_{t=x,y} (1 - e^{-\rho_t}) \end{cases} \quad (8)$$

其中, $\rho_x = \left(\frac{b_{cx}^{gt} - b_{cx}}{C_w}\right)^2$, $\rho_y = \left(\frac{b_{cy}^{gt} - b_{cy}}{C_h}\right)^2$, 如图 4(b) 所示。 C_w 和 C_h 分别是最小边界矩形的宽度和高度。当 α 接近 0 时, 距离损失的贡献减小, 反之亦然; 当 α 接近 $\frac{\pi}{4}$ 时, 距离损失的

贡献增加。形状损失计算式如(9)所示:

$$\Omega = \sum_{t=w,h} (1 - e^{-\omega_t})^\theta \quad (9)$$

其中, $\omega_w = \frac{|\omega - \omega^{gt}|}{\max(\omega, \omega^{gt})}$, $\omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})}$, 如图 4(c) 所示。 ω, h, ω^{gt} 和 h^{gt} 分别是预测框和真实框的宽度和高度, θ 控制形状损失的权重。最终损失表示为:

$$\mathcal{L}_{\text{box}} = 1 - \text{IoU} + \frac{\Delta + \Omega}{2} \quad (10)$$

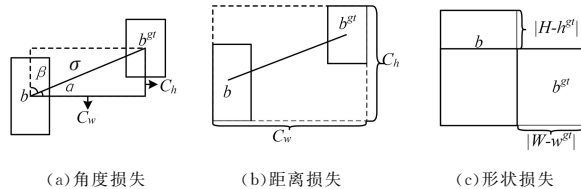


图 4 SIoU 损失函数

Fig. 4 SIoU loss function

4 实验

4.1 数据集

由于缺乏具有真实田间背景的合适的棉花叶片病害数据集, 本研究构建了一个常见棉花叶片病害的数据集作为研究对象。该数据集来自中国新疆乌鲁木齐的一个实验棉田。图像是使用苹果 iPhone 14 Pro 的后置摄像头拍摄的, 分辨率为 4032×3024 , 并以 jpg 格式保存。所有图像都是在自然光条件下获得的, 包括因天气和一天中的不同时间、树叶阴影和树叶重叠而导致的光线变化。该数据集包含本研究中 4 种典型棉花叶片病害 (轮纹病、褐斑病、叶霉病和白线病) 的 1566 张图像。部分实验数据集示例如图 5 所示, 其中第一行是原图, 第二行是注释图像。

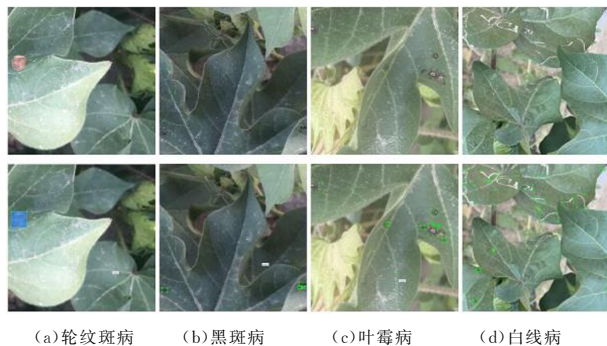


图 5 棉花病害数据集图像示例

Fig. 5 Image examples of cotton disease dataset

构建的数据集按 7:2:1 的比例分为训练集、验证集和测试集。训练集包含 1126 张图像, 验证集包含 282 张图像, 测试集包含 158 张图像。构建的数据集的具体信息如表 1 所列。

表 1 棉花病害数据集详情

病害	训练集	验证集	测试集	全部
轮纹斑病	286	72	40	398
黑斑病	421	106	59	586
叶霉病	227	57	32	316
白线病	192	49	27	288
全部	1126	282	158	1566

训练集从每个疾病数据集中随机选择,验证集用于在训练过程中确定模型的超参数,测试集用于评估模型的泛化能力。验证集和测试集之间的不兼容性确保了后续模型功能评估。

图像注释对于构建目标检测数据集至关重要。这项工作使用 LabelIMG 图像数据注释软件手动绘制棉花叶片病害病变的标记框。标记区域是病变周围最小的矩形框,包含尽可能少的背景区域。每个带注释的图像将生成一个 XML 文件,包括图片大小、标签名称和病变位置,如图 6 所示。



图 6 棉花病害数据集 XML 示例

Fig. 6 XML example of cotton disease dataset

4.2 实验指标

实验都是在 Nvidia RTX 3060 6GB 配置上进行的。具体参数设置如下:学习率设置为 0.01,学习率动量为 0.937,学习率衰减函数为余弦函数,模型优化器为随机梯度下降(SGD),优化器权重衰减因子为 0.0005,检测阈值为 0.4。批量大小设置为 8,训练进行了 300 个迭代周期,所有其他参数与 YOLOv5s 相同。这项工作使用了 3 个评估指标:精确度、召回率和 mAP。

4.3 对比实验

为了验证提出的方法的有效性,将其与数据集上的其他高级方法进行了比较,结果如表 2 所列。可以看出,所提方法分别将精度提高了 4.5 个百分点、2.1 个百分点、3.3 个百分点、3.7 个百分点和 6.3 个百分点。与基线模型相比,我们的方法将精度提高了 4.5 个百分点,达到了 90.4% 的精度,平均精度为 86.6%,召回率为 81.8%。实验证明,所提方法表现最佳,能在复杂场景中有效检测小目标和类似疾病。

表 2 对比实验

Table 2 Comparative experiments

模型	Precision	Recall	mAP
YOLOv5s	85.9	81.2	84.6
YOLO+SPD ^[24]	88.3	76.5	83.1
YOLO+SPC ^[29]	87.1	82	85.1
YOLO+lotus ^[27]	86.7	80.7	84.3
YOLOv8 ^[30]	84.1	79.9	83.8
YOLOv9 ^[31]	85.5	81.1	84.2
YOLO-GFM	90.4	81.8	86.6

4.4 消融实验

为了分析所提出方法的有效性,进行了一些消融研究。YOLOv5s 是基线模型,通过添加不同的模块来验证模型的性能。消融实验结果如表 3 所列。

表 3 消融实验

Table 3 Ablation experiment

YOLOv5s	GMA	FE	SiOU	Precision	Recall	mAP
T				85.9	81.2	84.6
T	T			87.8	81.7	84.7
T		T		87.6	80.5	85.2
T			T	87.9	80.9	84.2
T	T	T		88.5	81.0	85.5
T	T		T	88.5	79.9	84.8
T		T	T	88.2	80.6	84.6
T	T	T	T	90.4	81.8	86.6

结果表明,与原始模型相比,改进的 YOLO 模型在棉花叶片病害数据集上的检测性能显著提高。增强模型的 mAP、精确度和召回率分别提高了 2 个百分点、4.6 个百分点和 0.6 个百分点。这种情况表明,3 个模块的有效交互更适合在现实环境中检测棉花病害。

通过分析表 4 可以得出,添加所有模块后的高精度模型的精确度比基线模型高出了 4.5 个百分点,高精度模型的 mAP 比基线模型高出了 2.0 个百分点;高精度模型的 Parameters 比基线模型提高了 59.2%,高精度模型的 GFlops 比基线模型提高了 48.1%。GFM 通过增加 62.8% 的 Parameters 和 67.5% 的 GFLOPs,换取了 5.2% 的精确度提升和 2.4% 的 mAP 提升。这反映出模型性能提升与计算资源消耗之间存在非线性关系,说明算法改进的效率仍有优化空间。后续研究可考虑通过轻量化设计或效率优化策略来改善性能与资源消耗的平衡。

表 4 模型规模比较

Table 4 Comparison of model scales

Model	Parameters	GFLOPs	Precision/%	mAP/%
YOLOv5s	7.03×10^6	16.0	85.9	84.6
GFM	11.19×10^6	23.7	90.4	86.6

4.5 注意力比较

表 5 列出了在颈部网络的 C3 模块中添加不同的注意力模块的实验结果。与基线模型相比,引入这些注意力模块使精度提高了 0.2 个百分点、0.8 个百分点、1.2 个百分点、1.3 个百分点和 1.9 个百分点。结果表明,引入 GMA 显著提高了模型在复杂背景下对相似疾病的检测性能,因为 GMA 可以通过获得的上下文信息建立准确的全局关系,使模型能够学习更全面的对象特征并提高其判别能力,从而提高疾病检测的性能,也有利于在现实环境中检测棉花叶片疾病。

表 5 注意力模块比较

Table 5 Comparison of attention modules

模型	Precision	Recall	mAP
YOLOV5	85.9	81.2	84.6
SE ^[32]	86.1	81.7	84.7
CA ^[33]	86.7	80.7	84.3
ECA ^[34]	87.1	80.6	84.8
CBAM ^[35]	87.2	81.2	85.2
GMA	87.8	81.7	84.7

4.6 特征增强模块比较

为了验证 YOLO 模型中特征增强模块的有效性和必要性,将特征增强模块与 C3 进行了比较。如表 6 所列,结果表明 FE 模块的性能优于 C3。FE 模块提高了特征提取网络的

表达能力,使模型能够更好地捕捉疾病特征和背景特征之间的差异。

表 6 特征增强模块的比较

Table 6 Comparison of feature enhancement modules

(%)			
模型	Precision	Recall	mAP
YOLOV5	85.9	81.2	84.6
TK	87.6	80.5	85.2

4.7 损失函数比较

如表 7 所列,我们引入了不同的损失函数进行比较。与基线模型相比,引入不同损失函数后,精度分别提高了 1.4 个百分点、0.2 个百分点和 2.0 个百分点。SIoU 损失函数显著提高了模型在复杂背景下对类似疾病的检测性能。SIoU 通过更精确的边界框调整和几何特征捕获,提高了模型在病害检测中对小尺度病害点的准确性和识别能力,从而提高了病害检测

的性能,促进了棉花叶片病害在真实田间环境中的检测。

表 7 损失函数的比较

Table 7 Comparison of loss functions

(%)			
模型	Precision	Recall	mAP
CIoU ^[36]	85.9	81.2	84.6
EIoU ^[37]	87.3	79.5	84.6
WIoU ^[38]	86.1	79.8	93.7
SIoU	87.9	80.5	85.2

4.8 检测结果比较

图 7 表明,经过算法优化的检测模型在小目标识别性能上取得突破性进展。通过对比改进前后的检测结果可见,优化后的算法对小尺度目标的特征提取能力显著增强,检测灵敏度提升。在相同测试集上,改进模型的目标漏检率较原模型降低,特别是在复杂背景干扰下的微小目标检测稳定性得到有效改善。

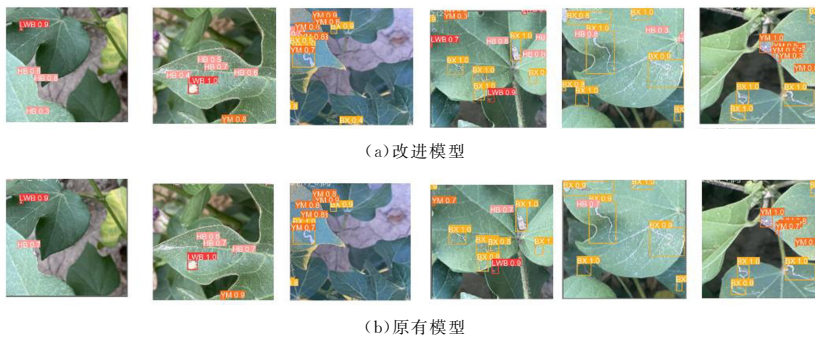


图 7 检测结果的对比

Fig. 7 Comparison of test results

结束语 本文提出了一种名为 YOLO-GFM 的棉花病害检测模型,该模型可以在真实的田间环境图像中准确检测 4 种棉花病害。FE 模块通过两个非线性激活和来自每个通道的特征信息来引导模型学习更多判别特征。GMA 模块通过获得的上下文信息建立准确的全局关系,捕获更多样化和丰富的特征,使模型能够学习更全面的对象特征并提高其判别能力。SIoU 损失函数显著提高了模型的回归性能。实验表明,该方法大大提高了 YOLO 在复杂真实环境中检测小目标和类似疾病的能力,为棉花病害的检测提供了强有力的技术支持,并为在类似复杂的真实环境中检测其他农业病害树立了先例。

参考文献

- [1] REDMON J. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
- [2] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//Computer Vision - ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, . Springer, 2016: 21-37.
- [3] LIN T. Focal loss for dense object detection [J]. arXiv:1708.02002, 2017.
- [4] TAN M, PANG R, LE Q V. EfficientDet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 10781-10790.
- [5] TIAN Z, SHEN C, CHEN H, et al. FCOS: Fully convolutional one-stage object detection[J]. arXiv:1904.01355, 2019.
- [6] LAW H, DENG J. CornerNet: Detecting objects as paired keypoints[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 734-750.
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 580-587.
- [8] CAI Z, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 6154-6162.
- [9] GIRSHICK R. Fast R-CNN[J]. arXiv:1504.08083, 2015.
- [10] REN S. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. arXiv:1506.01497, 2015.
- [11] HAQUE M A, MARWAHA S, DEB C K, et al. Deep learning-based approach for identification of diseases of maize crop[J]. Scientific Reports, 2022, 12(1): 6334.
- [12] ZHAO Y, SUN C, XU X, et al. RIC-Net: A plant disease classification model based on the fusion of inception and residual structure and embedded attention mechanism [J]. Computers and Electronics in Agriculture, 2022, 193: 106644.
- [13] FAN X, SUN T, CHAI X, et al. YOLO-WDNet: A lightweight and accurate model for weeds detection in cotton field[J]. Computers and Electronics in Agriculture, 2024, 225: 109317.
- [14] WANG B, HUA J, XIA L, et al. A defect detection method for

- Akidzuki pears based on computer vision and deep learning[J]. *Postharvest Biology and Technology*, 2024, 218:113157.
- [15] SYED-AB-RAHMAN S F, HESAMIAN M H, PRASAD M. Citrus disease detection and classification using end-to-end anchor-based deep learning model[J]. *Applied Intelligence*, 2022, 52(1):927-938.
- [16] JOHNSON J, SHARMA G, SRINIVASAN S, et al. Enhanced field-based detection of potato blight in complex backgrounds using deep learning[J]. *Plant Phenomics*, 2021, 2021:9835724.
- [17] BAO W, ZHU Z, HU G, et al. UAV remote sensing detection of tea leaf blight based on DDMA-YOLO[J]. *Computers and Electronics in Agriculture*, 2023, 205:107637.
- [18] TIAN Y, WANG S, LI E, et al. MD-YOLO: Multi-scale dense YOLO for small target pest detection[J]. *Computers and Electronics in Agriculture*, 2023, 213:108233.
- [19] FAN Z, LIU Q. Adaptive region-aware feature enhancement for object detection[J]. *Pattern Recognition*, 2022, 124:108437.
- [20] QI J, LIU X, LIU K, et al. An improved YOLOv5 model based on visual attention mechanism: Application to recognition of tomato virus disease[J]. *Computers and Electronics in Agriculture*, 2022, 194:106780.
- [21] ZHANG Y, MA B, HU Y, et al. Accurate cotton diseases and pests detection in complex background based on an improved YOLOX model[J]. *Computers and Electronics in Agriculture*, 2022, 203:107484.
- [22] GE C, DING X, TONG Z, et al. Advancing vision transformers with group-mix attention[J]. *arXiv: 2311. 15157*, 2023, 11: 113027-113041.
- [23] DENG C, WANG M, LIU L, et al. Extended feature pyramid network for small object detection[J]. *IEEE Transactions on Multimedia*, 2021, 24:1968-1979.
- [24] LI S, LIU C, TANG K, et al. Improved YOLOv5s algorithm for small target detection in UAV aerial photography[J]. *IEEE Access*, 2024.
- [25] JING R, ZHANG W, LIU Y, et al. An effective method for small object detection in low-resolution images[J]. *Engineering Applications of Artificial Intelligence*, 2024, 127:107206.
- [26] GAO J, GENG X, ZHANG Y, et al. Augmented weighted bidirectional feature pyramid network for marine object detection[J]. *Expert Systems with Applications*, 2024, 237:121688.
- [27] MA J, LU A, CHEN C, et al. YOLOv5-Lotus: An efficient object detection method for lotus seedpod in a natural environment[J]. *Computers and Electronics in Agriculture*, 2023, 206:107635.
- [28] SONG G, DU H, ZHANG X, et al. Small object detection in unmanned aerial vehicle images using multi-scale hybrid attention[J]. *Engineering Applications of Artificial Intelligence*, 2024, 128:107455.
- [29] HUANG M, YAN W, DAI W, et al. EST-YOLOv5s: SAR image aircraft target detection model based on improved YOLOv5s[J]. *IEEE Access*, 2023, 11:113027-113041.
- [30] VARGHESE R, SAMBATH M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness[C]// 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems(ADICS). IEEE, 2024:1-6.
- [31] GEVORGYAN Z. SiO loss: More powerful learning for bounding box regression[J]. *arXiv: 2205. 12740*, 2022.
- [32] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:7132-7141.
- [33] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 13713-13722.
- [34] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:11534-11542.
- [35] 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.
- [36] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]// Proceedings of the AAAI Conference on Artificial Intelligence. 2020: 12993-13000.
- [37] 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.
- [38] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: Bounding box regression loss with dynamic focusing mechanism[J]. *arXiv: 2301. 10051*, 2023.



WANG Hongqiang, born in 1997, post-graduate. His main research interest is image detection based on deep learning.



ZHAO Hui, born in 1972, Ph.D, professor, is member of CCF (No. 25440S). Her main research interests include artificial intelligence, affective computing, speech and digital image processing.