

基于 StyleGAN2 - ADA 和改进 YOLO v7 的葡萄 叶片早期病害检测方法

张林鎔^{1,2} 巴音塔娜¹ 曾庆松¹

(1. 新疆大学电气工程学院, 乌鲁木齐 830017; 2. 清华大学国家计算机集成制造系统工程技术研究中心, 北京 100084)

摘要: 为实现葡萄早期病害的快速准确识别, 针对葡萄病害的相似表型症状识别率低及小病斑检测困难的问题, 以葡萄黑腐病和黑麻疹病为研究对象, 提出了一种基于自适应鉴别器增强的样式生成对抗网络与改进的 YOLO v7 相结合的葡萄黑腐病和黑麻疹病的病斑检测方法。通过自适应鉴别器增强的样式生成对抗网络和拉普拉斯滤波器的方差扩充葡萄病害数据。采用 MSRCF 算法进行图像增强, 改善光照环境凸显病斑特征。以 YOLO v7 网络框架为基础, 将 BiFormer 注意力机制嵌入特征提取网络, 强化目标区域的关键特征; 采用 BiFPN 代替 PA - FPN, 更好地实现低层细节特征与高层语义信息融合, 以同时降低计算复杂度; 在 YOLO v7 的检测头部分嵌入 SPD 模块, 以提高模型对低分辨率图像的检测性能; 并采用 CIoU 与 NWD 损失函数组合对损失函数重新定义, 实现对小目标快速、准确识别。实验结果表明, 该方法病斑检测精确率达到 94.1%, 相比原始算法提升 5.7 个百分点, 与 Faster R - CNN、YOLO v3 - SPP 和 YOLO v5x 等模型相比分别提高 3.3、3.8、4.4 个百分点, 能够实现葡萄早期病害快速准确识别, 对于保障葡萄产业发展具有重要意义。

关键词: 葡萄; 病害识别; StyleGAN2 - ADA; 目标检测; 自注意力机制; YOLO v7

中图分类号: TP391; S24

文献标识码: A

文章编号: 1000-1298(2024)01-0241-12

OSID:



Grape Disease Detection Method Based on StyleGAN2 - ADA and Improved YOLO v7

ZHANG Linxuan^{1,2} BA Yintana¹ ZENG Qingsong¹

(1. School of Electrical Engineering, Xinjiang University, Urumqi 830017, China

2. National Computer Integrated Manufacturing System Engineering Research Center, Tsinghua University, Beijing 100084, China)

Abstract: Black rot and brown spot disease of grapes are diseases that seriously threaten grape yields, and identification of grape diseases early is of great significance for disease prevention and control and grape yield. However, current disease detection methods have a high leakage rate. The black rot and brown spot were taken as the research objects, a method for detecting grape black rot and brown spot based on adaptive discriminator enhanced style generation adversarial network combined with improved YOLO v7 was proposed. Firstly, the grape disease data were expanded by the adaptive discriminator enhanced style generation adversarial network + deblurring processing. Secondly, the MSRCF algorithm was used to enhance the image and improve the lighting environment to highlight the characteristics of disease spots. Finally, based on the YOLO v7 network framework, the BiFormer attention mechanism was embedded in the feature extraction network to strengthen the key features of the target area. BiFPN was used instead of PA - FPN to better realize multi-scale feature fusion and reduce computational complexity. SPD module was introduced in the detection head section of YOLO v7 to improve the detection performance of low-resolution images. The combination of CIoU and NWD loss function was used to redefine the loss function to achieve rapid and accurate identification of small targets. The experimental results showed that the accuracy of spot detection in this method reached 94.1%, which was 5.7 percentage points higher than that of the original algorithm, and 3.3 percentage points, 3.8 percentage points, and 4.4 percentage points higher than that of Faster R - CNN, YOLO v3 - SPP, and YOLO v5x models, respectively, which can realize the rapid and accurate identification of early grape diseases, which was of positive significance for ensuring the development of the grape industry.

Key words: grape; disease identification; StyleGAN2 - ADA; object detection; self-attention mechanism; YOLO v7

收稿日期: 2023 - 06 - 01 修回日期: 2023 - 08 - 01

基金项目: 新疆维吾尔自治区自然科学基金项目(2022D01C431)

作者简介: 张林鎔(1966—), 男, 教授, 博士, 主要从事智能制造技术与系统研究, E-mail: lxzhang@tsinghua.edu.cn

0 引言

我国是世界最大的葡萄生产国和消费国,葡萄作为一种重要的经济作物,广泛种植于全球许多地区^[1]。新疆作为葡萄栽培的重要地区之一,葡萄种植业在新疆具有重要的经济和社会意义。然而,许多病害引起了葡萄产量的严重降低,影响了葡萄园的经济效益。为了对葡萄病害进行有效的管理和控制,及早识别病害是很有必要的^[2]。

现有的葡萄病害识别手段通常是由经验丰富的专家进行甄别,不仅对专业知识要求高,并且需要耗费大量的时间和精力,对于表征相似的病害早期病害的识别不够准确,从而错过最佳防治时期。然而,随着计算机视觉^[3]、机器学习^[4]和人工智能等技术的不断发展^[5],计算机视觉成为现代智慧农业的重要组成部分^[6],实现了对农业生产过程中的各种信息的自动化获取和分析。当前基于计算机视觉在植物病害识别方面的研究主要分为两类,一类是以VGG16^[7]为代表的图像分类技术。彭红星等^[8]基于MobileNet V2模型引入双分支特征融合模块和坐标注意力机制识别葡萄病虫害,识别准确率达到89.16%。樊湘鹏等^[9]通过改进的VGG16模型引入迁移学习机制对葡萄的5类病害和健康叶片进行检测和分类,分类平均准确率达98.02%。何东健等^[10]提出一种改进的残差网络识别葡萄霜霉病,其识别准确率达到99.92%。PAYMODE等^[11]设计了一种通过卷积神经网络识别多作物病叶的深度学习模型,对葡萄和番茄的病叶和健康叶进行分类,其中葡萄病叶的分类准确率达到98.40%。YIN等^[12]提出了一种基于深度迁移学习和改进的MobileNetV3模型,在有限训练图像情况下提高了葡萄叶部病害的准确性,识别准确率达99.84%。以上研究表明,基于图像分类技术的葡萄病害识别已有一定程度的研究,且具有较高的识别准确率,但都是针对部分葡萄病害进行识别,无法对病斑区域进行精准定位。另一类是以Faster R-CNN^[13]、YOLO系列^[14]为代表的目标检测算法。CHEN等^[15]通过改进的YOLO v5目标检测网络对橡胶树病害进行识别,该模型对白粉病的检测精度为86.5%,对炭疽病的检测精度为86.8%。LIU等^[16]利用基于图像金字塔的多尺度特征检测、对象边界框维度聚类和多尺度训练对YOLO v3网络进行改进,用以检测番茄,其检测准确率达92.39%。以上研究论证了通过目标检测技术检测植物病虫害的可行性,但很少有研究关注葡萄病害的早期阶段,且存在大规模的目标疾病图像数据难以收集、现有模型检测精度

低、其改进方法不适用于早期小病斑检测等问题。

针对以上问题,本文以葡萄黑腐病和黑麻疹病为研究对象,提出一种基于自适应鉴别器增强的样式生成对抗网络(StyleGAN2-ADA)与改进的YOLO v7相结合的葡萄黑腐病和黑麻疹病的病斑检测方法。

1 研究方法

1.1 图像生成

生成对抗网络(GAN)是一种深度学习模型^[17],用于生成新的数据,如图像、音频和文本。GAN由生成器和判别器组成。生成器接收随机噪声作为输入,并生成与训练数据相似的虚假数据。判别器接收训练数据和生成器生成的数据。并对真实数据和虚假数据分类。生成器和判别器通过不断博弈,直到纳什均衡。因此,本文利用生成对抗网络对葡萄病害样本进行数据增产^[18]。

1.1.1 StyleGAN2-ADA

采用带有自适应鉴别器增强(Adaptive discriminator augmentation, ADA)的StyleGAN2来生成分辨率为256像素×256像素的葡萄黑腐病和黑麻疹病图像^[19]。在生成对抗网络中,当数据集很小时会导致判别器过拟合,从而产生对生成器的无效反馈,而StyleGAN2-ADA通过一种自适应判别器增强机制,有效地解决了数据限制的问题^[20]。由于本文所用数据样本较少,且无相似的大规模数据集,因此为使模型相对随机初始化参数具有更高的初始性能和更快的收敛速度,本文采用在FFHQ数据集上训练的预训练权重训练葡萄病害样本。StyleGAN2-ADA的生成器由映射网络和合成网络组成,判别器由特征提取器和分类器组成,结构如图1所示。

首先,将一组符合高斯分布的潜在向量 \mathbf{Z} 输入到由8层全连接层组成的映射网络中,将其编码为具有相同大小的中间向量 \mathbf{W} ,中间向量 \mathbf{W} 与输入向量 \mathbf{Z} 大小相同,用于控制不同的视觉特征,中间向量 \mathbf{W} 再通过仿射变换 A 生成控制风格的向量 \mathbf{S}_{style} 。合成网络的初始输入为 $4 \times 4 \times 512$,它由上采样层、卷积层、自适应实例规范化模块(AdaIN)以及风格控制向量 \mathbf{S}_{style} 及噪声 \mathbf{B} 组成。自适应实例规范化模块的表达式为

$$\text{AdaIN}(x_i, y) = y_{s,i} \frac{x_i - \mu(x_i)}{\sigma(x_i)} + y_{b,i} \quad (1)$$

式中 x_i ——第 i 层的特征图

$\text{AdaIN}(x_i, y)$ ——自适应实例规范化函数

$\mu(x_i)$ 、 $\sigma(x_i)$ ——各个通道特征图的方差和均值

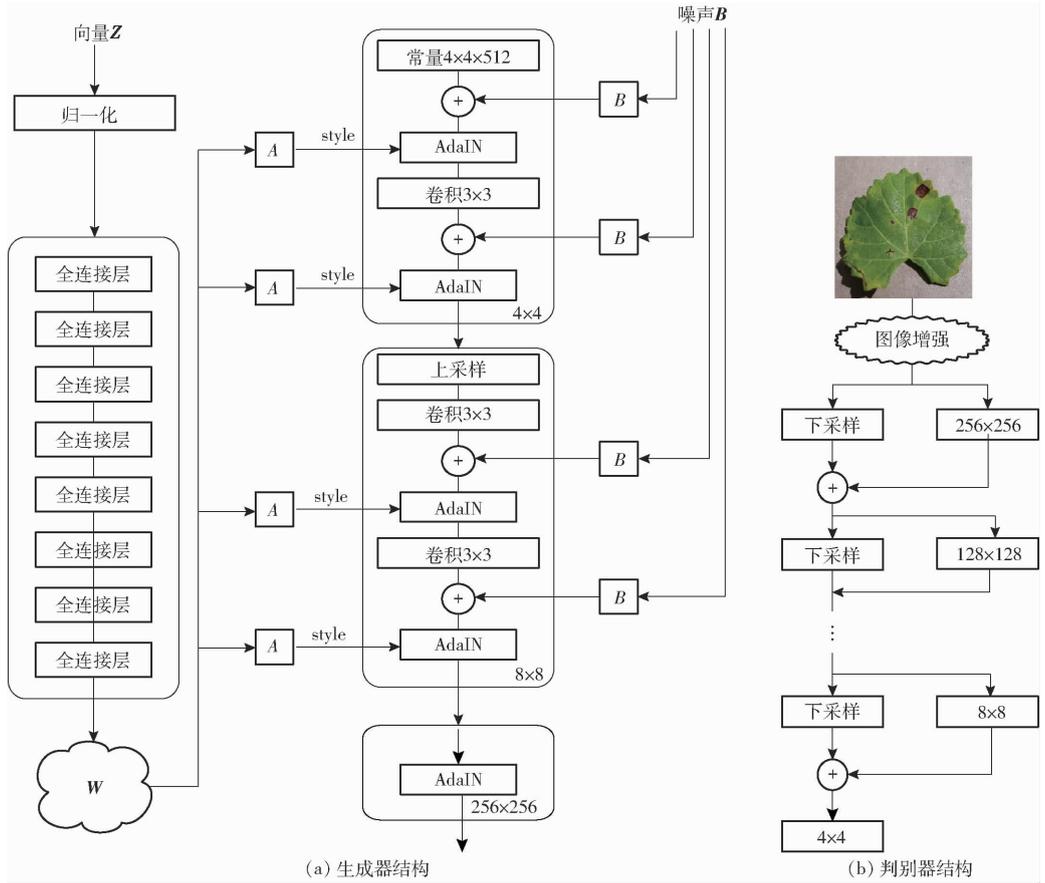


图 1 StyleGAN2 - ADA 生成器和判别器结构图

Fig. 1 StyleGAN2 - ADA generator and discriminator block diagram

$y_{s,i}, y_{b,i}$ ——收缩因子和偏差因子

StyleGAN2 - ADA 的自适应判别器增强流程图如图 2 所示。蓝色图像增强框代表与图像增强相关的操作,是 StyleGAN2 - ADA 中用于提高生成图像质量的一种技术,即设置一个随机增强概率 p ,根据过拟合的程度在不同的时间控制 p ,以 p 的概率决定是否对图像做图像增强,选定的概率 p 可以调整图像旋转角度、特定范围内调整色差。无论是真实图像还是生成图像都可以使用相同的增强方法进行图像增强。绿色的生成器 G 和判别器 D 代表正在训练中的网络。带有自适应鉴别器增强的 StyleGAN2,在损失函数和网络架构方面与 StyleGAN2 相似,但自适应鉴别器增强(ADA)有助于提高稳定性并避免过度拟合问题^[21]。因此,本文采用 StyleGAN2 - ADA 对葡萄病害数据进行增广。

1.1.2 去模糊处理

由于在生成图像中,存在一些模糊和扭曲图像,因此,采用拉普拉斯方法的方差进行过滤并丢弃一些生成的模糊图像。该方法通过传递图像的单个通道将其与 3×3 卷积核进行卷积来实现,连续二元函数 $f(x, y)$ 的拉普拉斯运算定义为

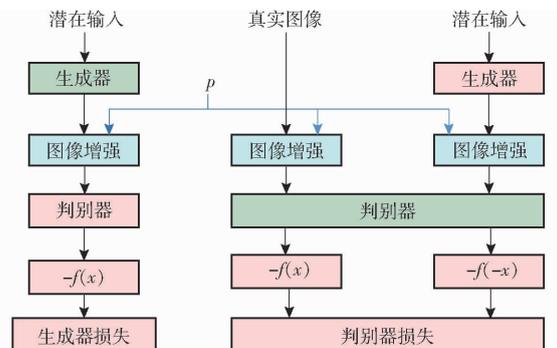


图 2 自适应判别器增强流程图

Fig. 2 Flowchart of adaptive discriminator augmentation

$$\nabla^2 f = \frac{\partial^2 f}{\partial^2 x} + \frac{\partial^2 f}{\partial^2 y} \quad (2)$$

x 的偏导数表示为

$$\frac{\partial^2 f}{\partial^2 x} = f(x+1, y) + f(x-1, y) - 2f(x, y) \quad (3)$$

y 的偏导数表示为

$$\frac{\partial^2 f}{\partial^2 y} = f(x, y+1) + f(x, y-1) - 2f(x, y) \quad (4)$$

结合 x 与 y 的偏导数为

$$\nabla^2 f = f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x, y) \quad (5)$$

形成 3×3 卷积核,用于对像素点 (x, y) 进行卷积:

$$\begin{pmatrix} (x-1,y-1) & (x,y-1) & (x+1,y-1) & 0 & 1 & 0 \\ (x-1,y) & (x,y) & (x,y+1) & \rightarrow & 1 & -4 & 1 \\ (x-1,y+1) & (x,y+1) & (x+1,y+1) & 0 & 1 & 0 \end{pmatrix}$$

将图像的单个通道与上述 3×3 卷积核进行卷积,同时获取相应的方差作为单幅图像的清晰度,通过排序选取最模糊的 10% 的图像认定为模糊图像,并选取拉普拉斯方差的极大值设为阈值,如果方差小于阈值,则图像被认为是模糊的,否则认为图像清晰^[22]。

1.2 具有色彩保护的多尺度 Retinex 算法

多尺度 Retinex 算法 (MSR) 是基于单尺度 Retinex 算法 (SSR) 开发的^[23]。MSR 使用线性加权改变图像中目标的特征,实现图像增强。本文采用由 Retinex 算法衍生出的 MSRCP 算法实现图像增强,改善光照环境,凸显叶片病斑特征。该方法采用增强图像与原始图像相结合的策略,将强度图像和原始图像的比例映射至每个通道,通过缩放因子 A 对 3 个通道进行色彩调整来得到增强后的图像,强度图像公式为

$$I_{nt} = (I_R + I_G + I_B) / 3 \quad (6)$$

式中 I_{nt} ——强度图像

I_R, I_G, I_B ——原始图像对应的三通道的值

反射因子计算公式为

$$A = \text{Min} \left(\frac{255}{B}, \frac{\text{Int}\alpha(i)}{\text{Int}(i)} \right) \quad (7)$$

其中 $B = \text{Max}(I_R(i) + I_G(i) + I_B(i)) \quad (8)$

式中 i ——像素索引

B ——像素索引 i 的通道最大值

$\text{Int}\alpha(i)$ ——经过 Retinex 算法处理后的单通道图像

$\text{Int}(i)$ ——经过 Retinex 算法处理后的强度图像

Max ——最大值函数

Min ——最小值函数

MSRCP 算法在恢复场景真实颜色的同时解决色偏的问题,从而最大限度地保留了物体原始特征,提高了图像质量^[24-25]。

1.3 目标检测模型及改进

YOLO v7 是一种单阶段目标检测技术^[26],YOLO v7 网络模型主要包括输入层、主干网络、特征融合、检测头 4 部分。主干网络骨干模块由若干个卷积层 (Conv)、批量归一化层 (BN) 和激活函数组成,用于提取不同尺度的图像特征,是特征提取的关键部分。特征融合部分结合了特征金字塔网络 (Feature pyramid network, FPN) 和路径聚合网络 (Path aggregation network, PANet) 的优点,形成 PA-

FPN 结构,使得不同层级的特征图进行高效融合。检测头部分通过生成 3 层不同大小的特征图,使用重参数化卷积 RepConv 输出预测结果。

1.3.1 加权双向特征金字塔网络

YOLO v7 网络在特征融合部分采用了 PA-FPN 结构^[27]对不同层级的特征图进行高效融合,PA-FPN 通过自上而下和自下而上的双路径实现底层特征信息和高层特征信息的融合。而加权双向特征金字塔网络 (BiFPN) 的主要思想是在不同的层级之间进行信息交互,每个层级都有多个分支,每个分支都有不同的特征尺度,通过多次堆叠来获得不同层级的语义特征^[28]。不同特征融合方式的结构图如图 3 所示。BiFPN 通过增加权重的方法区分融合过程中不同特征的重要程度,并且删除 PA-FPN 的冗余节点,能够在保持高精度的同时,具有较低的计算复杂度^[29]。因此,本文选用 BiFPN 代替了 PA-FPN,进一步加强特征融合性能以提高模型识别准确度。

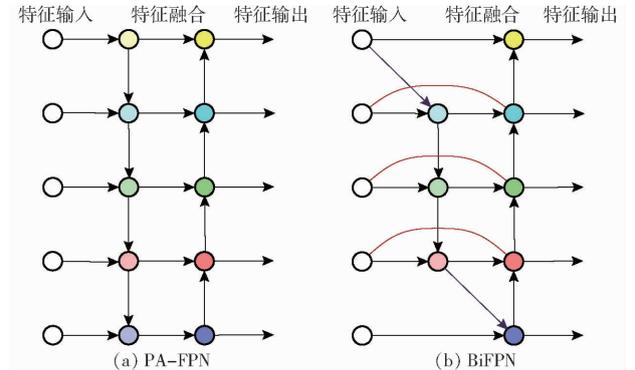


图 3 不同的特征融合结构

Fig. 3 Different feature fusion structures

1.3.2 BiFormer 注意力机制

CNN 由于感知域有限,需要多层堆叠来获取全局信息,但随着层数的增加,信息量减少,使得提取的特征的注意力集中在某些区域。而 Transformer^[30] 相对 CNN,其核心优势是采用了 self-attention 机制获取长距离上下文依赖,但其计算量大,占用内存大。而动态稀疏注意力机制 BiFormer,通过双层路由 (Bi-level routin) 实现灵活的计算分配和内容感知,将注意力模块的计算稀疏化,只计算一部分重要的注意力模块^[31]。本文引入 BiFormer 注意力机制的核心是将 attention 的计算分为 2 个阶段:

(1) 将图像分成多个粗粒度区域,在粗粒度区域上进行 self-attention,根据 Q 和 K 计算每 2 个粗粒度区域之间的相关性。从而得到关系矩阵,并对该关系矩阵稀疏化处理,只保留值最大的 K 个元素。

(2) 基于阶段 (1) 的粗粒度稀疏矩阵,进一步在细粒度区域上进行 self-attention 计算。

BiFormer 注意力机制的网络结构图如图 4 所示,输入图像 x 依次经过 4 个特征提取阶段,以获得不同大小的特征层。第 1 阶段由 N_i 个重叠块嵌入层 (Patch embedding) 和 BiFormer 模块构成,剩下 3 个阶段由 N_i 个块合并模块 (Patch merging) 和 BiFormer 模块构成。BiFormer 模块是由 3×3 的深度卷积、双层路由注意力 (Bi-level routing attention, BRA) 模块、归一化层 (Layer norm, LN) 和多层感知机 (Multi-layer perception, MLP) 组成。BRA 模块是

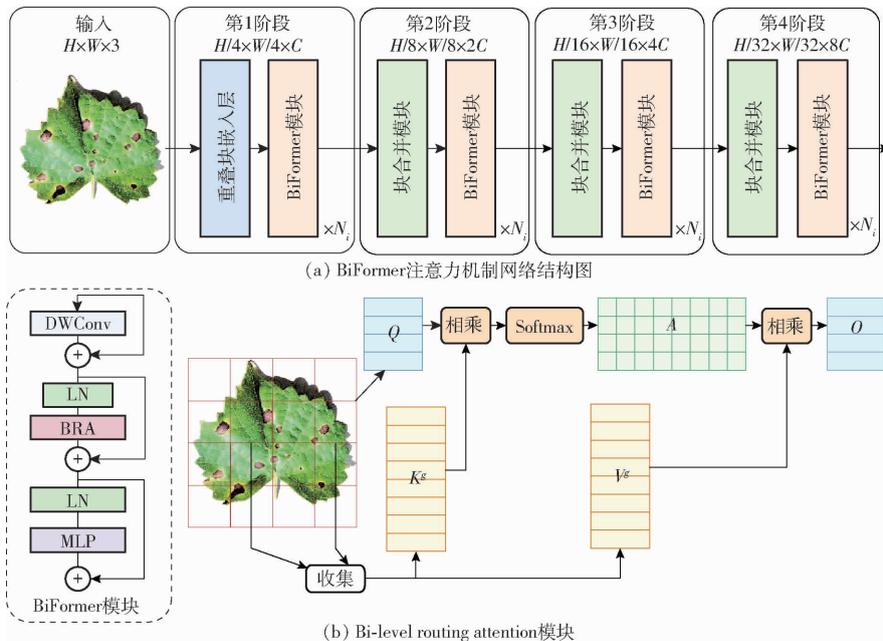


图 4 BiFormer 注意力机制架构

Fig. 4 Structure of BiFormer attention

1.3.3 NWD 损失函数

原始模型中的 CIoU 对微小目标的位置偏差较为敏感,降低了微小目标的检测性能,为提高网络对微小病斑的检测性能,本文引入 NWD (Normalized Wasserstein distance) 损失函数^[32]。采用 CIoU 损失与 NWD 损失函数组合的方式对边界框回归损失函数重新进行了定义,NWD 是以 Wasserstein 距离为基础,为微小物体设计的度量标准^[33],其过程分为 2 个阶段:

(1) 先将边界框建模为 2 维高斯分布

针对边界框 $A = (c_x, c_y, w, h)$,其中 (c_x, c_y) 、 w 、 h 分别表示中心坐标、边界框宽度和高度,建模为二维高斯分布 $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 。 $\boldsymbol{\mu}$ 为均值向量, $\boldsymbol{\Sigma}$ 为高斯分布的协方差矩阵。

$$\begin{cases} \boldsymbol{\mu} = \begin{bmatrix} c_x \\ c_y \end{bmatrix} \\ \boldsymbol{\Sigma} = \begin{bmatrix} \frac{w^2}{4} & 0 \\ 0 & \frac{h^2}{4} \end{bmatrix} \end{cases} \quad (9)$$

BiFormer 注意力机制的核心模块,首先将输入图像划分为 $S \times S$ 个相互不重叠的区域,再经过线性映射得到 Q, K, V 。其次针对每个区域定位该区域所关注的区域得到 Q' 和 K' ,通过 Q' 和 K' 的转置矩阵相乘得到不同区域之间亲和度图的邻接矩阵 A' ,邻接矩阵 A' 的元素表示两区域之间语义上的相关度。保留前 K 个关系最密切的区域得到路由索引矩阵 I' ,最后收集这 K 个区域的 K 和 V 得到 K^e ,合并进行 self-attention 计算。

(2) 计算 Wasserstein 距离定义并进行归一化
2 个二维高斯分布 $\mu_1 = N(\mathbf{m}_1 + \boldsymbol{\Sigma}_1)$ 和 $\mu_2 = N(\mathbf{m}_2 + \boldsymbol{\Sigma}_2)$ 之间的二阶 Wasserstein 距离定义为

$$W_2^2(\mu_1 - \mu_2) = \|\mathbf{m}_1 - \mathbf{m}_2\|_2^2 + \|\boldsymbol{\Sigma}_1^{\frac{1}{2}} - \boldsymbol{\Sigma}_2^{\frac{1}{2}}\|_F^2 \quad (10)$$

式中 $\|\cdot\|_F$ ——Frobenius 范数

W_2^2 ——距离度量

$\mathbf{m}_1, \mathbf{m}_2$ ——均值向量

$\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2$ ——协方差矩阵

因此对于边界框 A 和 B ,其中 $A = (cx_a, cy_a, w_a, h_a)$ 、 $B = (cx_b, cy_b, w_b, h_b)$,边界框的二维高斯分布 N_a 和 N_b 的二阶 Wasserstein 距离为

$$W_2^2(N_a, N_b) = \left\| \left(\left(cx_a, cy_a, \frac{w_a}{2}, \frac{h_a}{2} \right)^T, \left(cx_b, cy_b, \frac{w_b}{2}, \frac{h_b}{2} \right)^T \right) \right\|_2^2 \quad (11)$$

再使用它的指数形式归一化作为相似度度量,

即

$$NWD(N_a, N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a, N_b)}}{Q}\right) \quad (12)$$

式中 Q ——与数据集密切相关的常数

$NWD(N_a, N_b)$ ——变量间的正态分布水平距离

1.3.4 空间到深度卷积层模块

对于图像分辨率较低的目标检测任务,传统 CNN 因为使用跨步卷积及池化层导致丢失一些细粒度信息,所以其检测性能较低^[34]。因此,为提升低分辨率图像的检测效果,在 YOLO v7 的检测头部引入了一种新的卷积神经网络结构 SPD-Conv^[35]。其整体结构如图 5 所示。

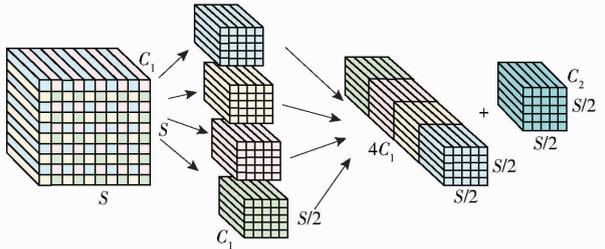


图 5 SPD-Conv 模块整体架构

Fig.5 Overall framework of SPD-Conv

SPD-Conv 使用 SPD 层和 Non-strided 层来代替传统 CNN 模块中的池化层和跨步式卷积。SPD 层将特征图 $X(S, S, C_1)$ 以 S_e 为单位划分为一系列子特征图

$$f_{0,0} = X[0:S:s_{scale}, 0:S:s_{scale}] \quad (13)$$

$$f_{1,0} = X[1:S:s_{scale}, 1:S:s_{scale}] \quad (14)$$

$$f_{s_{scale}-1, s_{scale}-1} = X[s_{scale}-1:S:s_{scale}, s_{scale}-1:S:s_{scale}] \quad (15)$$

式中 S ——特征图边长 C_1 ——特征图深度

s_{scale} ——图像比例

$f_{0,0}$ 、 $f_{1,0}$ 、 $f_{s_{scale}-1, s_{scale}-1}$ ——子特征图

通过划分得到 $\frac{xy}{s_{scale}^2}$ 个子特征图,再将这些子特征图沿着通道维度拼接,获得中间特征图 $X'(\frac{S}{s_{scale}}, \frac{S}{s_{scale}}, s_{scale}^2 C_1)$ 。再通过 Non-strided 层,利用增加的卷积层中的可学习参数来减少通道数,减少信息的无差别损失,尽可能保留所有的判别特征信息,将中间特征图 $X'(\frac{S}{s_{scale}}, \frac{S}{s_{scale}}, s_{scale}^2 C_1)$ 转化为最终特征图 $X'(\frac{S}{s_{scale}}, \frac{S}{s_{scale}}, C_1)$ 。SPD-Conv 模块在针对低分辨率图像和极小物体检测任务时表现良好,大大减少了信息丢失。

1.3.5 整体结构

YOLO v7 模型是当前较新的单阶段目标检测模型,相较以往模型具有更快的卷积操作,能够检测更多细粒度对象。然而当待检测任务多为小目标、分辨率较低时,YOLO v7 的检测精确度欠佳。因此,本

研究针对葡萄病斑数据,将从以下 4 方面对 YOLO v7 进行改进:

(1)在特征提取后添加 BiFormer 注意力机制,使模型忽略无关信息,关注重要信息,有效提高模型在复杂背景下的特征提取能力。

(2)将原模型的特征融合方式(PA-FPN)替换为加权双向特征金字塔网络(BiFPN),进一步加强了不同层结构之间的特征融合,同时减少特征冗余问题,降低计算复杂度。

(3)在 YOLO v7 的检测头部分引入 SPD-Conv 模块,以提高网络对低分辨率图像的检测效果。

(4)引入以 Wasserstein 距离为基础,采用 CIoU 与 NWD 损失函数组合对损失函数重新进行定义,提高网络对小目标的检测效果。

基于以上思路,改进 YOLO v7 完整的网络层结构如图 6 所示。

2 实验

2.1 实验环境搭建

模型开发语言为 Python,实验过程在 Windows 10 操作系统下进行,处理器型号为 i7-13700K,显卡型号为 RTX3080,编程平台为 Pycharm,深度学习框架为 PyTorch 1.11。网络模型训练参数设置中,学习率为 0.01,图像批量大小为 8,训练轮数为 250。

2.2 评价指标

评价指标主要包括图像生成评价指标及目标检测评价指标。为判定图像生成质量,采用 FID (Fréchet inception distance) 分数作为评价指标。针对目标检测模型,采用精确率 (Precision)、召回率 (Recall)、平均精度均值 (Mean average precision, mAP) 及每秒传输帧数 (Frame per second, FPS) 作为评价指标。

2.3 数据集

选用 Plantvillage 植物病害数据集作为训练集,该数据集包括苹果、葡萄、玉米、西红柿等 14 种作物的 26 种农作物病害样本信息^[36]。通过筛选和整理,本文选取病害早期表征相似的黑腐病图像 725 幅及黑麻疹病图像 700 幅作为研究对象,并通过对抗生成网络对该样本进行数据增广,数据集样本示例如图 7 所示。

3 实验结果与分析

3.1 图像生成结果及分析

为了提高检测精度,均衡样本数量,增强目标检测网络的鲁棒性,本文引入 DCGAN、WGAN、

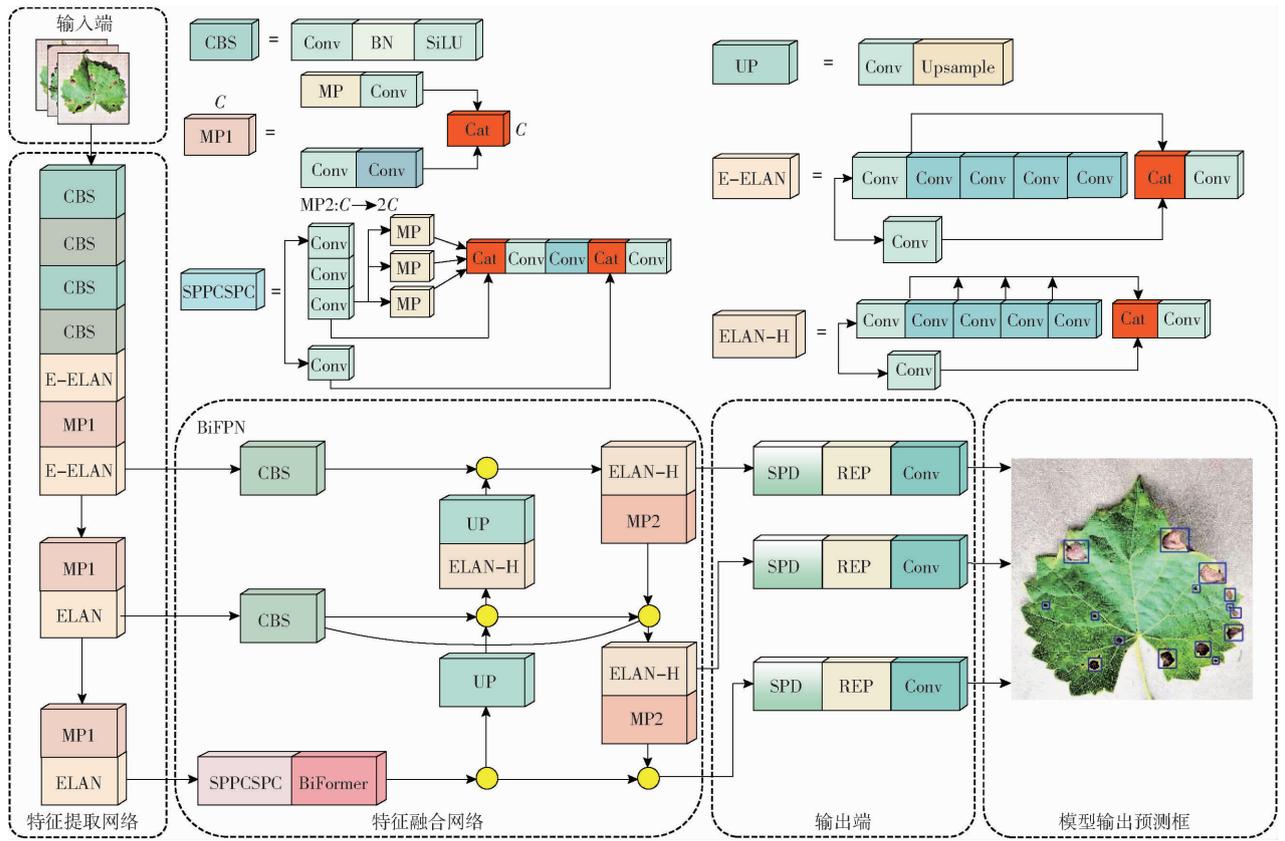


图 6 改进 YOLO v7 网络框图

Fig. 6 Improved YOLO v7 network architecture

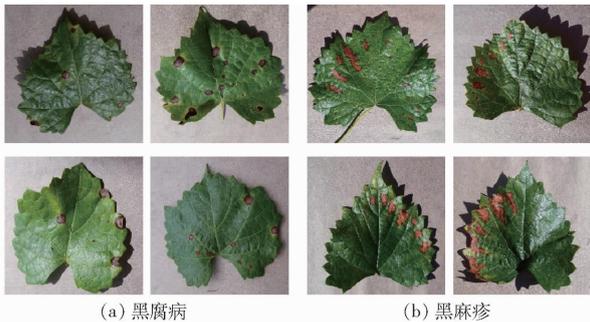


图 7 葡萄病害部分数据集

Fig. 7 Grape disease partial datasets

StyleGAN2、StyleGAN2 - ADA + 去模糊处理对葡萄黑腐病和黑麻疹病数据进行数据扩充。为客观评价生成图像的质量,采用 FID 分数判定生成图像的质量。FID 分数是在同一区域内真实图像与生成图像的 Inception 特征向量之间的距离,用于评估不同对抗生成模型的性能,在稳健性和效率方面表现良好。当真实图像和生成图像的特征越相近则 FID 分数越小,生成图像的质量就越好。FID 分数计算表达式为

$$FID = \| m_r - m_g \|^2 + T_r (C_r + C_g - 2 (C_r C_g)^{\frac{1}{2}}) \quad (16)$$

式中 m_r ——真实图像特征均值

C_r ——真实图像协方差矩阵

m_g ——生成图像特征均值

C_g ——生成图像协方差矩阵

T_r ——矩阵的迹

不同数据增强方式的对比实验结果如表 1 所示,通过比较 DCGAN、WGAN、StyleGAN2、StyleGAN2 - ADA 的 FID 分数验证不同模型生成葡萄病叶的性能,FID 分数越低表明生成图像越接近真实样本,模型性能越好。由表 1 可知,通过 DCGAN 和 WGAN 生成的葡萄病害图像 FID 分数较高,生成图像质量较低。通过 StyleGAN 生成的葡萄病害图像相比前两者 FID 分数显著降低,但生成的图像仍存在较多伪影及边界不清晰的图像。而 StyleGAN2 - ADA 生成的葡萄病害图像 FID 分数为 30.95,在 4 种模型中最低,且生成图像质量高、特征清晰。因此本文采取 StyleGAN2 - ADA 和拉普拉斯滤波器的方差对葡萄病叶数据进行扩充,为展示 StyleGAN2 - ADA 模型的生成效果,选取训练不同时期的生成图像进行对比分析。

表 1 不同数据增强方式的 FID 分数

Tab. 1 FID scores for different ways to augment data

模型	DCGAN	WGAN	StyleGAN2	StyleGAN2-ADA
生成图像				
FID 分数	299.50	241.30	88.86	30.95

由图8可知,随着迭代周期的增加,生成的图像质量越来越好,生成图像与真实图像之间的FID分数越来越低,生成图像质量也不断提高。迭代周期为40时模型生成的图像质量较低,迭代周期为80时,生成图像已初具雏形但图像质量仍较低。迭代周期为140时生成图像已无法通过肉眼分辨真假。

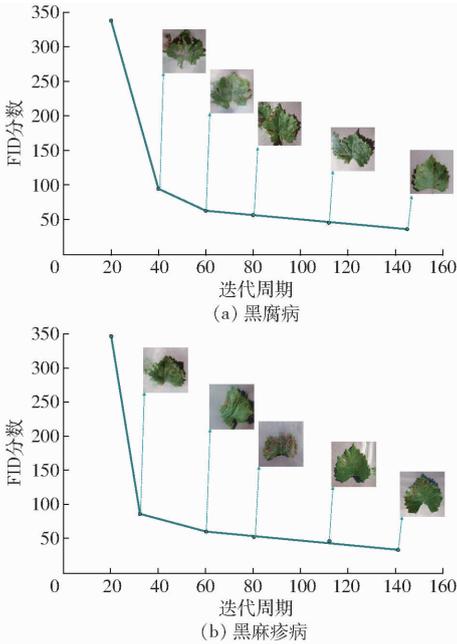


图8 训练过程的FID分数

Fig. 8 FID score of training process

以上实验表明,StyleGAN2-ADA模型能够通过真实图像的概率分布,以判别器和生成器相互对抗的方式生成符合条件的图像,解决了样本数量不足的问题。

3.2 数据集标注及分析

根据数据集结构,本文按照病害种类对数据样本进行二分类,并采用LabelImg图像标注软件对病斑叶片进行人工标注。按照比例8:2划分训练集和测试集,抽取部分标注文件进行可视化分析,结果如图9所示。

由图9可知,目标框大小不均匀,但比值大多分布在0.02~0.2之间,属于小目标偏多,不易检测。

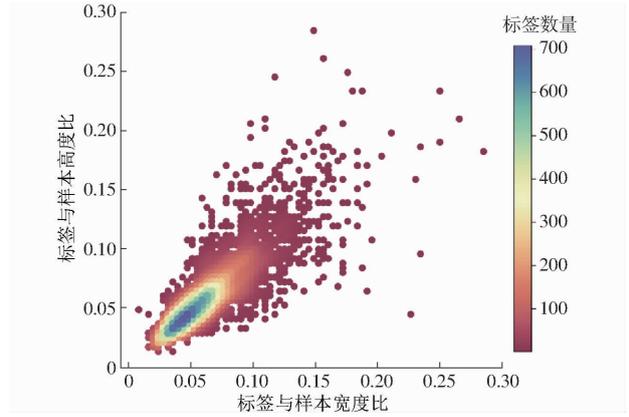


图9 训练集文件可视化

Fig. 9 Visualization of training set files

针对上述问题,本文通过对单阶段目标检测模型YOLO v7进行改进,设计出能够满足葡萄叶片小病斑检测的模型。

3.3 病斑检测结果及分析

本文设计了3组对比实验,所有实验均在同一实验平台进行。第1组是YOLO v7模型和所有改进后的模型进行对比。第2组是将改进后的模型与其他YOLO系列模型以及目前主流的目标检测算法对比。第3组是分析注意力机制的影响。

3.3.1 病斑检测模型消融实验

为验证各个改进模块对葡萄病斑检测模型的影响,本文在YOLO v7的基础上设计消融实验。通过StyleGAN2-ADA+去模糊处理生成新的样本与原始数据混合并作为新数据集训练的实验组记为YOLO v7-A。将所有图像数据经过MSRCP处理后再训练得到的网络模型记为YOLO v7-B。在YOLO v7网络模型中加入BiFormer注意力机制得到的模型记为YOLO v7-C。在YOLOv7-C实验组中采用BiFPN结构代替PA-FPN结构得到的网络模型记为YOLO v7-D。在YOLO v7-D实验组中加入SPD模块得到的网络模型记为YOLO v7-E。在YOLO v7-E实验组的基础上采用CIoU与NWD损失函数的组合对损失函数重新进行定义得到的网络模型为本文方法。消融实验结果如表2所示。

表2 消融实验结果

Tab. 2 Ablation experiment

方法	数据扩充	MSRCP	BiFormer	BiFPN	SPD	NWD	准确率/%	召回率/%	mAP _{0.5} /%	FPS/(f·s ⁻¹)
YOLO v7							88.4	85.5	89.3	78.12
YOLO v7-A	★						89.8	86.5	92.6	71.42
YOLO v7-B	★	★					90.3	88.8	93.6	71.42
YOLO v7-C	★	★	★				92.0	90.6	94.1	72.46
YOLO v7-D	★	★	★	★			93.5	89.9	94.9	80.65
YOLO v7-E	★	★	★	★	★		93.8	88.6	94.6	78.85
本文方法	★	★	★	★	★	★	94.1	87.7	95.4	77.12

注:★表示添加该模块。

通过 YOLO v7 与 YOLO v7 - A 的结果对比可知, 利用 StyleGAN2 - ADA 和拉普拉斯滤波器的方差扩充数据后, 模型准确率提高 1.4 个百分点, 证明了本文采用的数据扩充方法的有效性。通过 YOLO v7 与 YOLO v7 - B 的结果对比可知, 采用 MSRCPP 算法进行图像增强能有效提高网络的识别准确率。通过 YOLO v7 与 YOLO v7 - C 的结果对比可知, 嵌入 BiFormer 注意力机制后, 模型准确率提高 1.8 个百分点, 网络对病斑的关注度提高。通过 YOLO v7 与 YOLO v7 - D 的结果对比可知, 采用 BiFPN 结构进行特征融合后, 模型准确率提高 1.5 个百分点, 且检测速度高于基线模型, 证明该特征融合机制能够更好地实现低层细节特征与高层语义信息融合, 同时减少特征冗余问题, 降低了计算复杂度。通过 YOLO v7 与 YOLO v7 - E 结果对比可知, SPD 模块能够提高模型准确率和 $mAP_{0.5}$, 且仍然满足快速识别的要求。通过 YOLO v7 与本文方法的结果对比可知, 将边界框回归损失定义为 CIOU 损失与 NWD 损失函数组合的形式能够有效降低检测损失, 提高检测准确率, 损失函数曲线如图 10 所示。通过本文方法的结果可知, 最终改进 YOLO v7 模型的准确率达到 94.1%, $mAP_{0.5}$ 达到 95.4%, 模型检测速度虽不及基线模型, 但下降幅度低, 仍然满足实时检测要求, 能够在病斑检测精度和速度之间有很好的平衡, 验证了改进模型的可行性。模型训练过程的准确率曲线如图 11 所示。为直观地比较改进模型与基线模

型的检测效果, 图 12 展示了测试集人工标注图像和改进前后的检测结果。

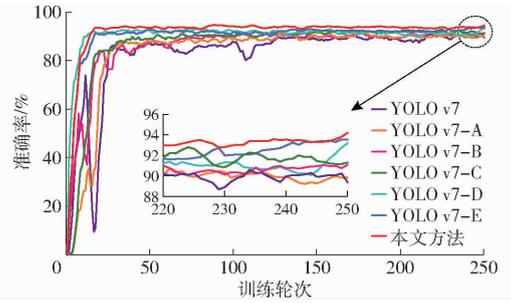


图 11 训练过程的准确率曲线

Fig. 11 Precision of training process

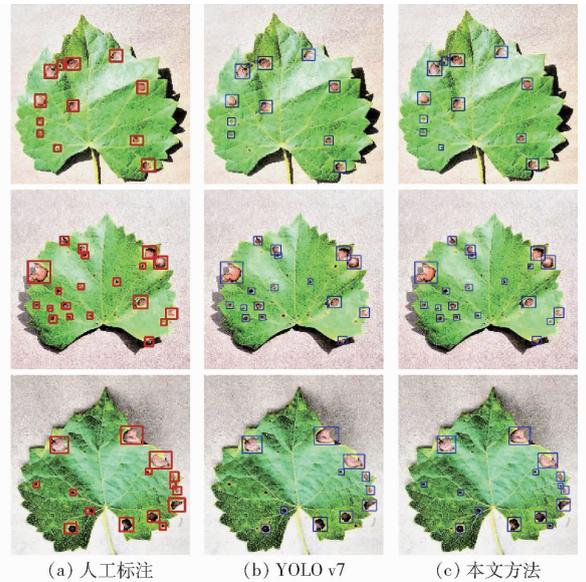


图 12 模型检测效果图

Fig. 12 Model detection renderings

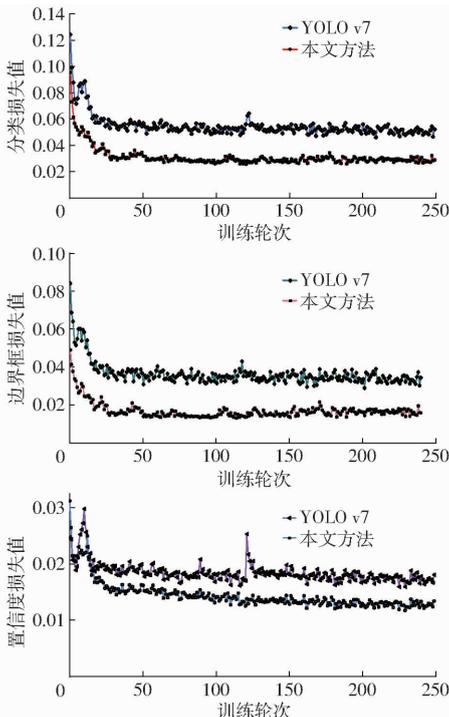


图 10 不同模型损失函数曲线

Fig. 10 Loss function of different models

3.3.2 病斑检测模型对比实验

为了验证网络性能, 选取两阶段目标检测模型 (Faster R - CNN)、单阶段目标检测算法 (SSD、YOLO v5x、YOLO v3 - SPP、YOLO v5 + BiFPN) 以及文献[37]中的玉米叶片病害检测模型 (CTR_YOLO v5n)、文献[38]中的苹果叶片病害检测模型 (BTC - YOLO v5s) 进行对比, 结果如表 3 所示。

表 3 不同模型葡萄病害数据集检测效果

Tab. 3 Detection results of grape disease in different models

模型	准确率/%	$mAP_{0.5}/%$	FPS/($f \cdot s^{-1}$)
Faster R - CNN	90.8	93.7	21.32
SSD	83.5	87.1	59.51
YOLO v3 - SPP	90.3	92.3	73.52
YOLO v5x	89.7	93.6	63.29
YOLO v5 + BiFPN	90.6	92.5	75.12
CTR_YOLO v5n	92.2	93.9	156.25
BTC - YOLO v5s	91.7	93.8	99.01
本文方法	94.1	95.4	77.12

从表3可以看出,本文方法的准确率相较两阶段目标检测模型 Faster R-CNN 高 3.3 个百分点,相较单阶段目标检测算法 SSD、YOLO v3-SPP、YOLO v5x、YOLO v5 + BiFPN、CTR_YOLO v5n、BTC-YOLO v5s 分别高 10.6、3.8、4.4、3.5、1.9、2.4 个百分点。两阶段目标检测网络 Faster R-CNN 召回率和 $mAP_{0.5}$ 较高,但模型推理速度最慢。单阶段目标检测模型 SSD 由于其单一的特征融合方式检测准确率最低。YOLO v5x 和 YOLO v5 + BiFPN 召回率较高,但准确率和 $mAP_{0.5}$ 较低。文献[37]中的玉米叶片病害检测模型 CTR_YOLO v5n 和文献[38]中的苹果叶片病害检测模型 BTC-YOLO v5s 由于使用有较低网络深度和宽度的 YOLO v5s 和 YOLO v5n 具有较快的检测速度,但准确率和 $mAP_{0.5}$ 低于本文所提算法。综合结果表明:本文算法在病斑检测的检测精度和检测速度之间具有较好的平衡,能够更好地对葡萄病害的病斑进行准确检测。

3.3.3 注意力机制对比实验

为检验不同位置下添加 BiFormer 注意力机制对检测效果的影响,分别在特征提取网络的不同层添加 BiFormer 注意力机制,实验结果如表4所示。

表4 注意力机制位置对结果的影响

Tab.4 Influence of attention mechanisms at different locations %

注意力机制	引入位置	准确率	召回率	$mAP_{0.5}$
		90.2	88.8	93.6
BiFormer	Layer 为 24	90.2	87.4	94.1
	Layer 为 37	90.8	88.0	93.5
	Layer 为 50	91.2	90.1	93.7
	SPPCSPC 后	92.0	90.6	94.1

通过表4可看出,在不同位置引入 BiFormer 注意力机制,准确率和 $mAP_{0.5}$ 有所变化,且随着层数增加,提升效果更显著,而在 SPPCSPC 结构后传输的特征将在特征融合结构中传播,因此在 SPPCSPC 结构后采取微小改进,对后续特征融合也会产生很大影响,因此本文在 SPPCSPC 后引入 BiFormer 注意力机制,提高对于病斑目标的表达能力。

为验证 BiFormer 注意力机制的有效性,分别与 SE 注意力机制、CBAM 注意力机制、SimAM 注意力机制对比,实验结果如表5所示。此外,还利用 Grad-CAM 算法^[39]展示了原始模型和添加不同注意力模块后模型的热力图,以可视化网络在检测过程中关注的重点区域,结果如图13所示。

根据热力图图像可以看出,YOLO v7 模型仅能关注部分病斑目标,其他注意力机制在关注病斑目标的同时也同时关注了部分背景目标,而本文所用

表5 不同注意力机制的检测结果

Tab.5 Detection results for different attention

模型	mechanisms		
	准确率	召回率	$mAP_{0.5}$
YOLO v7	90.2	88.8	93.6
YOLO v7 + SimAM	90.9	90.7	92.1
YOLO v7 + SE	88.2	78.8	87.3
YOLO v7 + CBAM	90.4	84.1	92.3
YOLO v7 + BiFormer	92.0	90.6	94.1

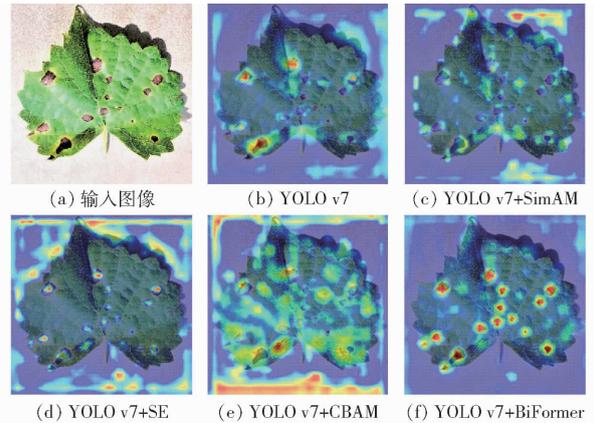


图13 Grad-CAM 可视化结果

Fig.13 Grad-CAM visualization results

的 BiFormer 注意力机制能够有效强化目标区域的关键特征,更好的关注葡萄叶片的病斑区域,从而有效减少训练损失。

4 结论

(1)采用葡萄病害数据,针对葡萄黑腐病和黑麻疹病数据,提出了利用自适应鉴别器增强的样式生成对抗网络和拉普拉斯滤波器的方差对葡萄黑腐病及黑麻疹病数据扩充,采用具有色彩保护的多尺度的 Retinex 算法 (MSRCP) 进行图像增强,通过 Labellmg 工具对葡萄叶片病斑进行标注。

(2)提出一种改进的 YOLO v7 目标检测模型,针对该数据集分辨率较低,且目标多为小目标的特点,引入 SPD 和针对微小目标的度量 NWD 损失函数。在 SPPCSPC 结构后引入 BiFormer 注意力机制,强化目标区域的关键特征。采用 BiFPN 代替了 PA-FPN,以实现更高层的特征融合。改进后的目标检测模型准确率为 94.1%, $mAP_{0.5}$ 达到 95.4%,相较原始模型分别提升 5.7 个百分点和 6.1 个百分点。

(3)所提方法能够快速准确地检测出叶片不同种类的病斑,通过检测病斑的形态、大小、颜色、分布等特征初步判断葡萄病害的严重程度,并及时采取相应的防治措施,对葡萄园病害管理和控制具有积极意义。

参 考 文 献

- [1] FRAIWAN M, FAOURI E, KHASAWNEH N. Multiclass classification of grape diseases using deep artificial intelligence[J]. *Agriculture*, 2022, 12(10): 1542.
- [2] YIN X, LI W, LI Z, et al. Recognition of grape leaf diseases using MobileNetV3 and deep transfer learning[J]. *International Journal of Agricultural and Biological Engineering*, 2022, 15(3): 184-194.
- [3] 王利伟, 徐晓辉, 苏彦莽, 等. 基于计算机视觉的葡萄叶部病害识别研究[J]. *江苏农业科学*, 2017, 45(23): 222-225.
WANG Liwei, XU Xiaohui, SU Yanmang, et al. Study on recognition of grape leaf diseases based on computer vision[J]. *Jiangsu Agricultural Sciences*, 2017, 45(23): 222-225. (in Chinese)
- [4] 韩文霆, 崔家伟, 崔欣, 等. 基于特征优选与机器学习的农田土壤含盐量估算研究[J]. *农业机械学报*, 2023, 54(3): 328-337.
HAN Wenting, CUI Jiawei, CUI Xin, et al. Estimation of farmland soil salinity content based on feature optimization and machine learning algorithms[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2023, 54(3): 328-337. (in Chinese)
- [5] 刘成良, 贡亮, 苑进, 等. 农业机器人关键技术研究现状与发展趋势[J]. *农业机械学报*, 2022, 53(7): 1-22, 55.
LIU Chengliang, GONG Liang, YUAN Jin, et al. Current status and development trends of agricultural robots[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(7): 1-22, 55. (in Chinese)
- [6] ZHANG F, CHEN Z, ALI S, et al. Multi-class detection of cherry tomatoes using improved Yolov4-tiny model [J]. *International Journal of Agricultural and Biological Engineering*, 2023, 16(2): 225-231.
- [7] HU G, WU H, ZHANG Y, et al. A low shot learning method for tea leaf's disease identification[J]. *Computers and Electronics in Agriculture*, 2019, 163: 104852.
- [8] 彭红星, 徐慧明, 刘华甯. 融合双分支特征和注意力机制的葡萄病虫害识别模型[J]. *农业工程学报*, 2022, 38(11): 156-165.
PENG Hongxing, XU Huiming, LIU Huanai. Model for identifying grape pests and diseases based on two-branch feature fusion and attention mechanism[J]. *Transactions of the CSAE*, 2022, 38(11): 156-165. (in Chinese)
- [9] 樊湘鹏, 许燕, 周建平, 等. 基于迁移学习和改进 CNN 的葡萄叶部病害检测系统[J]. *农业工程学报*, 2021, 37(6): 151-159.
FAN Xiangpeng, XU Yan, ZHOU Jianping, et al. Detection system for grape leaf diseases based on transfer learning and updated CNN[J]. *Transactions of the CSAE*, 2021, 37(6): 151-159. (in Chinese)
- [10] 何东健, 王鹏, 牛童, 等. 基于改进残差网络的田间葡萄霜霉病病害程度分级模型[J]. *农业机械学报*, 2022, 53(1): 235-243.
HE Dongjian, WANG Peng, NIU Tong, et al. Classification model of grape downy mildew disease degree in field based on improved residual network[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, 53(1): 235-243. (in Chinese)
- [11] PAYMODE A S, MALODE V B. Transfer learning for multi-crop leaf disease image classification using convolutional neural network VGG[J]. *Artificial Intelligence in Agriculture*, 2022, 6: 23-33.
- [12] YIN X, LI W, LI Z, et al. Recognition of grape leaf diseases using MobileNetV3 and deep transfer learning[J]. *International Journal of Agricultural and Biological Engineering*, 2022, 15(3): 184-194.
- [13] HU C, SHI Z, WEI H, et al. Automatic detection of pecan fruits based on Faster RCNN with FPN in orchard [J]. *International Journal of Agricultural and Biological Engineering*, 2022, 15(6): 189-196.
- [14] PENG H, XUE C, SHAO Y, et al. Litchi detection in the field using an improved YOLOv3 model[J]. *International Journal of Agricultural and Biological Engineering*, 2022, 15(2): 211-220.
- [15] CHEN Z, WU R, LIN Y, et al. Plant disease recognition model based on improved YOLOv5[J]. *Agronomy*, 2022, 12(2): 365.
- [16] LIU J, WANG X. Tomato diseases and pests detection based on improved Yolo V3 convolutional neural network[J]. *Frontiers in Plant Science*, 2020, 11: 898.
- [17] 邓博, 贺春林, 徐黎明, 等. 生成对抗网络文字生成图像算法综述[J]. *计算机工程与应用*, 2022, 58(23): 42-55.
DENG Bo, HE Chunlin, XU Liming, et al. Text-to-Image synthesis: survey of state-of-the-art[J]. *Computer Engineering and Applications*, 2022, 58(23): 42-55. (in Chinese)
- [18] 袁培森, 吴茂盛, 翟肇裕, 等. 基于 GAN 网络的菌菇表型数据生成研究[J]. *农业机械学报*, 2019, 50(12): 231-239.
YUAN Peisen, WU Maosheng, ZHAI Zhaoyu, et al. Mushroom phenotypic generation based on generative adversarial network [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2019, 50(12): 231-239. (in Chinese)
- [19] KARRAS T, AITTALA M, HELLSTEN J, et al. Training generative adversarial networks with limited data[J]. *Advances in Neural Information Processing Systems*, 2020, 33: 12104-12114.
- [20] WOODLAND M K, WOOD J, ANDERSON B M, et al. Evaluating the performance of StyleGAN2-ADA on medical images [C]// *Simulation and Synthesis in Medical Imaging: 7th International Workshop*, 2022: 142-153.

- [21] SITU Z, TENG S, LIU H, et al. Automated sewer defects detection using style-based generative adversarial networks and fine-tuned well-known CNN classifier[J]. IEEE Access, 2021, 9: 59498 – 59507.
- [22] BANSAL R, RAJ G, CHOUDHURY T. Blur image detection using Laplacian operator and Open-CV[C]//2016 International Conference System Modeling & Advancement in Research Trends (SMART). IEEE, 2016: 63 – 67.
- [23] WEI C, WANG W, YANG W, et al. Deep retinex decomposition for low-light enhancement[J]. arXiv Preprint, arXiv:1808.04560, 2018.
- [24] 司永胜, 肖坚星, 刘刚, 等. 基于 MSRCF 与改进 YOLO v4 的躺卧奶牛个体识别方法[J]. 农业机械学报, 2023, 54(1): 243 – 250, 262.
SI Yongsheng, XIAO Jianxing, LIU Gang, et al. Individual identification method of lying cows based on MSRCF and improved YOLO v4 model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 243 – 250, 262. (in Chinese)
- [25] 衡宝川, 肖迪, 张翔. 结合 MSRCF 增强的夜间彩色图像拼接算法[J]. 计算机工程与设计, 2019, 40(11): 3200 – 3204, 3211.
HENG Baochuan, XIAO Di, ZHANG Xiang. Night color image mosaic algorithm combined with MSRCF[J]. Computer Engineering and Design, 2019, 40(11): 3200 – 3204, 3211. (in Chinese)
- [26] WANG C Y, BOCHKOVSKIY A, LIAO H. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J]. arXiv e-prints, arXiv: 2207. 02696, 2022.
- [27] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759 – 8768.
- [28] 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.
- [29] GUO Y, CHEN S, ZHAN R, et al. SAR ship detection based on YOLOv5 using CBAM and BiFPN[C]//IGARSS 2022—2022 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2022: 2147 – 2150.
- [30] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. arXiv: 1706.03762, 2017.
- [31] ZHU L, WANG X, KE Z, et al. BiFormer: vision transformer with bi-level routing attention[J]. arXiv Preprint, arXiv:2303.08810, 2023.
- [32] ZHOU X, JIANG L, GUAN X, et al. Infrared small target detection algorithm with complex background based on YOLO – NWD[C]//Proceedings of the 4th International Conference on Image Processing and Machine Vision, 2022: 6 – 12.
- [33] WANG J, XU C, YANG W, et al. A normalized Gaussian wasserstein distance for tiny object detection[J]. arXiv Preprint, arXiv: 2110.13389, 2021.
- [34] LI A, SUN S, ZHANG Z, et al. A multi-scale traffic object detection algorithm for road scenes based on improved YOLOv5 [J]. Electronics, 2023, 12(4): 878.
- [35] SUNKARA R, LUO T. No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects[C]//Machine Learning and Knowledge Discovery in Databases, 2023: 443 – 459.
- [36] 李书琴, 陈聪, 朱彤, 等. 基于轻量级残差网络的植物叶片病害识别[J]. 农业机械学报, 2022, 53(3): 243 – 250.
LI Shuqin, CHEN Cong, ZHU Tong, et al. Plant leaf disease identification based on lightweight residual network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(3): 243 – 250. (in Chinese)
- [37] MA L, YU Q, YU H, et al. Maize leaf disease identification based on YOLOv5n algorithm incorporating attention mechanism [J]. Agronomy, 2023, 13(2): 521.
- [38] LI H, SHI L, FANG S, et al. Real-time detection of apple leaf diseases in natural scenes based on YOLOv5[J]. Agriculture, 2023, 13(4): 878.
- [39] 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.