

基于混合扩张卷积和注意力的黄瓜病害严重度估算方法

李凯雨¹ 朱昕怡¹ 马浚诚² 张领先¹

(1. 中国农业大学信息与电气工程学院, 北京 100083; 2. 中国农业科学院农业环境与可持续发展研究所, 北京 100081)

摘要: 自动和准确地估计病害的严重度对病害管理和产量损失预测至关重要。针对传统病害严重度估算步骤复杂且低效,难以实现在田间场景下精准估算问题,提出了一种基于混合扩张卷积和注意力机制改进 UNet (Mixed dilated convolution and attention mechanism optimized UNet, MA-UNet) 的病害严重度估算方法。首先,针对病斑尺寸不一、形状不规则问题,提出混合扩张卷积块 (Mixed dilation convolution block, MDCB) 增加感受野并保持病斑信息的连续性,提升病斑分割精度。其次,为了克服复杂背景的影响,利用注意力机制 (Attention mechanism) 对空间维度和通道维度进行相关性建模,获得每个像素类内响应和通道间的依赖关系,缓解背景对网络学习带来的影响。最后,计算病害分割图中病斑像素与叶片像素的比率来获得严重度。基于田间条件下收集的黄瓜霜霉病和白粉病图像进行了验证,并与全卷积网络 (Fully convolutional network, FCN)、SegNet、UNet、PSPNet、FPN、DeepLabV3+ 进行比较。结果表明,MA-UNet 优于比较方法,能够满足复杂环境下健康叶片和病斑的分割需求,平均交并比为 84.97%,频权交并比为 93.95%。基于 MA-UNet 分割结果估计黄瓜叶部病害严重度的决定系数为 0.965 4,均方根误差为 1.083 7%。该研究可为人工智能在农业中快速估计和控制病害严重度提供参考。

关键词: 黄瓜病害; 病害严重度; 扩张卷积; 注意力机制; 语义分割

中图分类号: TP391

文献标识码: A

文章编号: 1000-1298(2023)02-0231-09

OSID:



Estimation Method of Leaf Disease Severity of Cucumber Based on Mixed Dilated Convolution and Attention Mechanism

LI Kaiyu¹ ZHU Xinyi¹ MA Juncheng² ZHANG Lingxian¹

(1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. Institute of Environment and Sustainable Development in Agriculture,
Chinese Academy of Agricultural Sciences, Beijing 100081, China)

Abstract: Automatic and accurate estimation of disease severity is critical for disease management and yield loss prediction. Traditional disease severity estimation steps are complicated and inefficient, which makes it challenging to achieve accurate estimation in practical scenarios. A disease severity estimation method was proposed based on mixed dilated convolution and attention mechanism to improve UNet (MA-UNet). Firstly, to solve the problem of different sizes and irregular shapes of lesions, the mixed dilation convolution block (MDCB) was proposed to increase the receptive field and maintain the continuity of lesion information to improve the accuracy of lesion segmentation. Secondly, to overcome the influence of complex background, the attention mechanism (AM) was used to model the correlation between the spatial and channel dimensions. It can obtain the response within each pixel class and the dependency between channels to alleviate the background's influence on network learning. Finally, the ratio of diseased lesion pixels to leaf pixels in the disease segmentation map was calculated to obtain the severity. It was validated based on cucumber downy mildew and powdery mildew images collected under field conditions and compared with fully convolutional network (FCN), SegNet, UNet, PSPNet, FPN, and DeepLabV3+. MA-UNet can meet the segmentation requirements of leaves and lesions in complex environments, with a mean intersection over union 84.97% and a value of frequency-weighted intersection over union value of 93.95%. Moreover, it can accurately estimate the severity of cucumber

收稿日期: 2022-03-15 修回日期: 2022-05-09

基金项目: 国家自然科学基金项目 (62176261)

作者简介: 李凯雨 (1994—), 女, 博士生, 主要从事图像处理 and 农业信息技术研究, E-mail: 1162103154@qq.com

通信作者: 张领先 (1970—), 男, 教授, 博士生导师, 主要从事农业信息技术研究, E-mail: zhanglx@cau.edu.cn

leaf diseases, the correlation coefficient was 0.965 4, and the RMSE was 1.083 7%. The results showed that MA - UNet outperformed the comparison methods in refining lesion segmentation and accurately estimating disease severity. The research result can provide a reference for artificial intelligence to estimate and control disease severity in agriculture rapidly.

Key words: cucumber disease; disease severity; dilation convolution; attention mechanism; semantic segmentation

0 引言

黄瓜是世界上第三大消费蔬菜,其年产量超过 $8.3 \times 10^7 \text{ t}^{[1]}$ 。霜霉病和白粉病等病害是导致黄瓜品质下降和农民经济损失的主要原因之一。病害严重度的准确估算是精准用药和预测作物损失的关键^[2],因此,估计病害严重度非常必要。

视觉估计是量化病害严重度的常规方法,即人眼对可见光范围内感知的症状严重程度进行赋值。比例尺是视觉估计中最准确的工具,人们利用比例尺测量植物上病害的数量和面积进行估计^[3],是视觉估计中最准确的工具^[4]。然而,由于不同病害症状的异质性和视觉估计的主观性,视觉估计并不可靠。标准面积图(Standard area diagrams, SAD)可以用来提高严重度估计的准确性^[5],但是 SAD 只能在白色或黑色等简单的背景下估计特定病害的严重度。

基于数字图像分析的严重度评估准确、可重复,且这种方法有成熟的处理过程。首先,通过图像预处理或手动操作去除背景噪声。其次,将颜色转换与数学形态学操作、阈值处理等相结合分割出病斑。最后,计算病斑面积和叶片面积的比例得到病害严重度。这些分割算法依赖于手工设计的图像特征,在受控条件下简单高效。但是,田间条件下采集的病害图像含有大量的土壤背景和复杂光照等噪声,且不同种类病害及不同严重程度的病害特征具有多样性和复杂性,手工设计的特征往往是低层次的特征,对光照变化、复杂背景和不同病斑位置很敏感,因此上述方法难以适应田间条件下的病害严重度估算。

深度学习(Deep learning, DL)受益于其自动学习图像的低级特征和高级语义特征,已经成为图像处理最先进的技术^[6]。将病害严重度转换为计算机视觉问题时,基于 DL 的病害严重度估计可分为 3 类,即基于分类的方法、基于回归的方法和基于语义分割的方法。基于分类的方法采用卷积神经网络(Convolutional neural network, CNN),通过将严重度分为严重程度的类别或区间将其转换为一个分类问题^[7-8],如健康、早期、中期和晚期。基于分类的严重度估算取得了较高的准确率^[9],但是此种分类方

法在田间试验中并不利于评估处理的有效性。基于回归的方法和基于语义分割的方法可以获得严重度的百分比值,很有参考价值^[3]。张领先等^[10]构建一个 CNN 模型估计黄瓜霜霉病的严重度,以手动去除背景的病害图像作为输入,证明了 CNN 的准确性优于浅层机器学习模型,决定系数 R^2 达到 0.919 0。基于语义分割的方法是通过深度神经网络代替人为设计特征的分割方法获得病害分割图。在简单背景下取得了较为准确的估计结果^[11-13],但自然条件下拍摄的图像不可避免地存在杂乱的背景,会降低严重度估计的准确性^[14]。鉴于此,WANG 等^[14]设计两阶段的病害分割模型以减少复杂背景的影响,首先使用 DeepLabV3 + 分割病害叶片,对分割出的病害叶片使用 UNet 分割出病斑。然而,两阶段的分割消耗更多的计算资源,增加严重度估计的复杂性。因此,有必要探索一种更高效且适合自然条件下的黄瓜叶部病害图像严重度估计方法。

本文以田间条件下采集的黄瓜霜霉病和白粉病为研究对象,提出一种基于混合扩张卷积和注意力机制改进 UNet(MA - UNet)的病害严重度估算方法。首先选取 UNet 为病害分割基础模型,添加批归一化层有助于模型训练;其次,设计具有不同膨胀率的混合扩张卷积块,捕捉不同尺寸的病斑特征,提升分割精度;同时,利用注意力机制有效增强网络的语义信息和位置信息,缓解背景对网络学习的影响。最后,计算病害分割图中病斑像素与叶片像素的比率来获得严重度,并与 FCN、SegNet、UNet、FPN、PSPNet 和 DeepLabV3 + 进行比较,以验证本文方法的有效性。

1 数据来源

1.1 图像采集

黄瓜霜霉病和白粉病图像采集于天津市农业科学院农业创新基地的 5 号日光温室。采集时要求每幅图像中保留一片完整的黄瓜叶片,背景中带有土壤或其他叶片等噪声。使用尼康 Coolpix S3100 型数码相机在自动模式下共采集 153 幅图像,其中霜霉病图像 76 幅,白粉病图像 77 幅。在采集过程中没有使用光学变焦或闪光灯。原始图像尺寸并不适用于模型,使用图像缩放方法降低计算成本,提高计

算效率,将图像尺寸统一调整为 224 像素 × 224 像素。预处理后病害图像样例如图 1 所示。

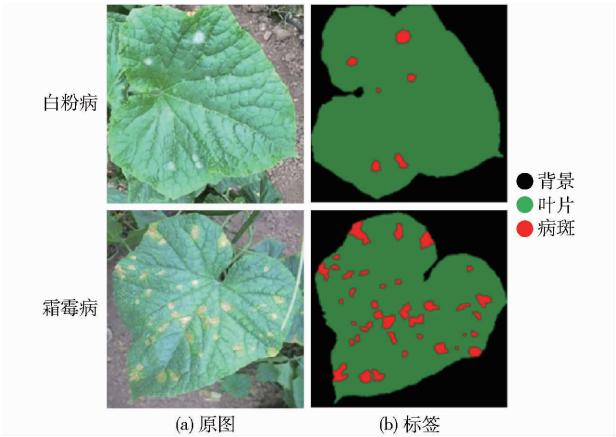


图 1 统一图像尺寸后的病害图像及标注样例

Fig. 1 Disease image and annotated sample after unifying image size

1.2 图像标注

使用 Matlab Image Labeler App (MathWorks Inc. , 美国) 对病害图像进行了像素级的标注。在标注过程中,图像中可能有不止一片叶片,背景中还包含其他不完整的叶片。因此,叶片和病斑类别的标记只在中心完整叶片上进行。背景、叶片和病斑的类别分别被标记为 0、1 和 2。为了便于可视化,病害标记样例如图 1 所示。标记之后,病害严重度 S 计算式为

$$S = \frac{P_{\text{lesion}}}{P_{\text{lesion}} + P_{\text{leaf}}} \times 100\% \tag{1}$$

式中 P_{lesion} ——分割图中病斑像素点数量
 P_{leaf} ——分割图中健康叶片像素点数量

1.3 图像增强

根据分层抽样法,将采集的 153 幅黄瓜叶部病害图像随机划分为训练集、验证集和测试集,其比例为 6:2:2。由于训练集中只有 93 幅图像,为了丰富数据的多样性以及适应卷积神经网络所需要的大量数据,对采集的每种病害数据集进行数据增强以防止过拟合。增强的策略主要是水平翻转、垂直翻转、随机缩放、顺时针旋转和逆时针旋转等方式。扩增后,训练数据集中的图像数量增加到 2 976 幅,扩充示例如图 2 所示。验证集和测试集的图像数量均为 30 幅,不做扩充处理。

2 黄瓜叶部病害分割及严重度估算模型建立

对采集到的黄瓜病害叶片图像严重度估计流程如图 3 所示。通过训练集训练提出的 MA - UNet 分割模型,验证集用来调整模型的超参数并初步评估模型的性能。其次,在测试集上对分割模型的性能

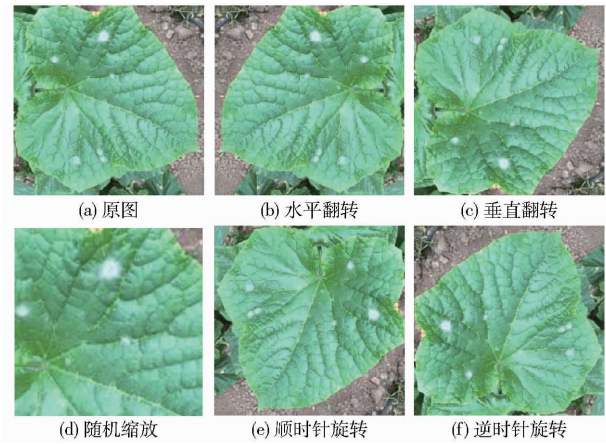


图 2 训练样本增强示例

Fig. 2 Examples of training data augmentation

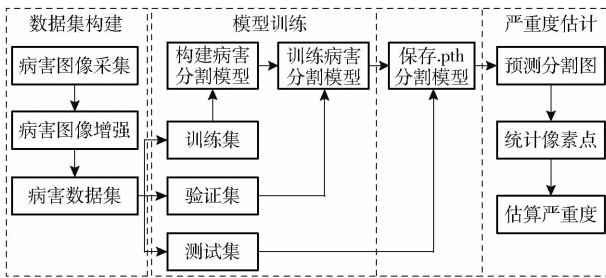


图 3 严重度估算整体流程图

Fig. 3 Overall flow chart of severity estimation

进行最终对比,并得到分割图。最后,根据分割图分别计算健康叶片和病斑的像素数,并根据式(1)计算病害严重度。

2.1 模型构建

2.1.1 分割网络

由于 UNet 架构在解决图像到图像映射任务方面的有效性^[11],本文选择其作为构建模型的基础。最初为医学图像分割任务提出的 UNet 模型采用了自编码器结构,编码器从输入图像中提取特征,解码器进行每个像素分类来重构分割的输出。另外,编、解码器之间的跳跃连接能为输出图像精确定位。UNet 在医学图像分割^[15]、作物分割^[16-17]等任务上取得了可喜的成果。本文在原始 UNet 的基础上,在每个卷积层后引入批归一化层 (Batch normalization, BN),实现对输入数据的标准正态分布来保证层间数据分布的稳定性,能够有效提升网络学习能力,加快网络收敛速度。在编码器与解码器的中间引入具有不同膨胀率的混合扩张卷积块 (Mixed dilation convolution block, MDCB),减少下采样导致图像细节的丢失。其次,在解码器里添加注意力机制,优化网络提取病害图像的位置信息和语义信息,进一步提升分割精度。本文提出的分割网络如图 4 所示。

2.1.2 混合扩张卷积块

在多种语义分割任务中,深度卷积神经网络已

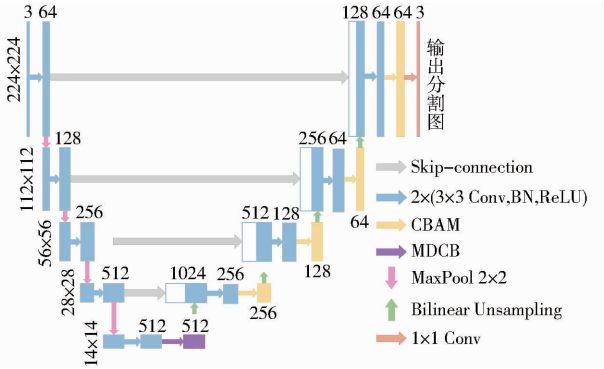


图4 MA-UNet网络结构

Fig. 4 Block diagram of MA-UNet for disease image segmentation

被证明可以有效地提取图像的特征表示,然而,传统卷积神经网络中的池化层会导致图像细节丢失。扩张卷积可以克服这个限制^[18]。扩张卷积引入膨胀率,可以在不丢失空间信息的情况下以指数方式增加感受野。

但是扩张卷积在卷积核中的2个像素之间填充零,导致该核的感受野仅覆盖具有棋盘图案的区域。因此,仅对非零值的位置进行采样,从而丢失了一些相邻信息,容易形成“网格化”。受 JIANG 等^[19]研究启发,本文提出一种混合扩张卷积块 (MDCB) 来避免网格效应,有效地捕获高分辨率特征图并保留更多空间信息,如图 5 所示。混合扩张卷积块由 3 个带有不同膨胀率的扩张卷积块组成,扩张卷积块由 4 个卷积层、批归一化层以及 LReLU (Leaky ReLU) 激活函数组成。其中膨胀率遵循类似锯齿波的启发式方法,即许多层组合在一起形成膨胀率增加的“上升边缘”,文中使用的膨胀率 r 依次为 1、2 和 5。大膨胀率的卷积可以提取和生成更多大尺度对象的抽象特征,而小膨胀率的卷积对于小目标更好。通过结合不同膨胀率的扩张卷积, MDCB 块能够提取具有尺寸不一、形状不规则的病斑特征,对病害分割更有效。

2.1.3 注意力机制

注意力模型有很多种,卷积注意力机制

(Convolutional block attention module, CBAM)^[20] 是一种简单有效的注意力模块,沿着通道和空间 2 个独立维度依次进行推理,2 个维度的注意力相辅相成,最后与输入特征进行自适应特征优化,结构如图 6 所示。此外, CBAM 模块的一大优势是轻量级,可无缝集成到任何神经网络中,无需增加计算开销。

通道注意力针对输入特征 $F(H \times W \times C)$,其中, H 为图像高, W 为图像宽, C 为通道数) 分别进行空间全局平均池化和最大池化,得到 2 个 $1 \times 1 \times C$ 特征,通过 ReLU 激活函数将它们分别送到 1 个 2 层神经网络,对全局特征进行非线性学习,得到 2 个特征并相加,再通过 Sigmoid 激活函数得到权重系数 M_c 。最后,将权重系数 M_c 与原始特征 F 相乘,得到通道维度上缩放后的新特征 F' 。与通道注意力类似,空间注意力针对特征 $F'(H \times W \times C)$,首先在通道维度上进行平均池化和最大池化,聚合特征映射的信息压缩得到 2 个 $H \times W \times 1$ 的特征,并拼接在一起。然后,通过卷积操作激活其空间特征值,得到缩放后的空间特征图 M_s 。在病害严重度估算研究中, 本文探索了注意力的位置及数量,实验表明将 CBAM 添加到解码器中,提取解码器中的细节特征,并与编码器的特征相融合,有益于模型提取病害图像的病斑特征,抑制背景的影响。

2.2 损失函数

病害数据集中背景类别、病斑类别和健康叶片类别的像素数量显然不均衡。这 3 个类别之间的差异可能会对训练产生不平衡的影响,而忽略了病斑像素的重要性。因此,实验中使用了加权交叉熵损失函数来减少这种影响。每个类别的权重 W_m 根据 BADRINARAYANAN 等^[21] 中位频率平衡 (Median frequency balance) 计算。最终用于数据集的背景、健康叶片和病斑的权重分别为 1.000 0、0.228 6 和 3.453 2。

$$W_m = \frac{A_f}{f_m}$$

(2)

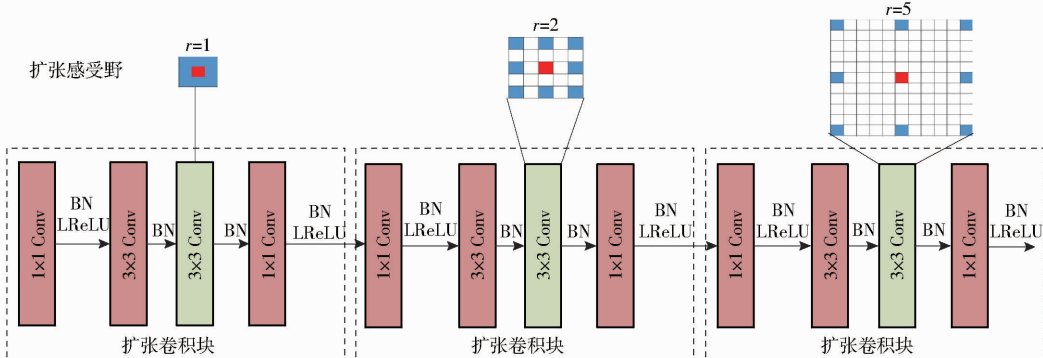


图5 MDCB模块结构示意图

Fig. 5 Illustration of MDCB block

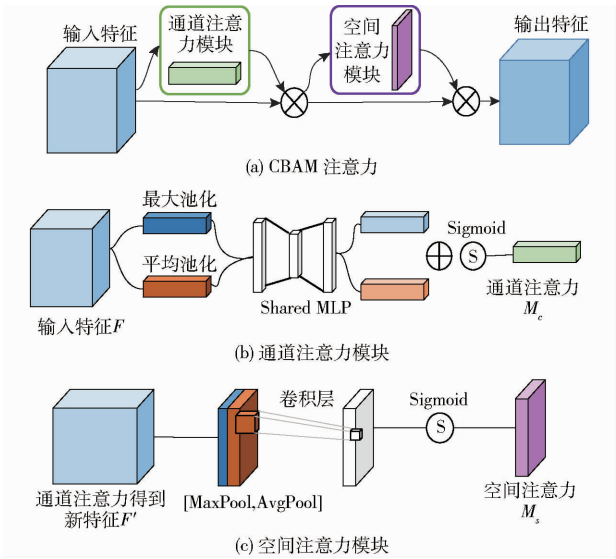


图 6 CBAM 注意力模块

Fig. 6 CBAM attention block

式中 f_m —— m 类像素出现的频率
 A_f ——类频率中值
应用类别权重之后的像素级交叉熵损失函数 L 为

$$L = -\frac{1}{N} \sum_{n=1}^N \sum_{m=0}^M W_m y_{n,m} \ln p_{n,m} \quad (3)$$

式中 N ——总样本数
 M ——目标类别数量,取 2
 $y_{n,m}$ ——符号函数,如果第 n 个样本属于第 m 类取 1,否则为 0
 $p_{n,m}$ ——第 n 个样本预测为第 m 类的概率

2.3 实验设置

本文中的模型是基于 PyTorch 的深度学习库实现的,并使用 NVIDIA Quadro P2000 GPU(5GB) 进行训练。训练过程中使用动量为 0.9 的随机梯度下降法(Stochastic gradient descent, SGD)。在设备允许的情况下,多次调整参数,最终初始学习率设置为 0.007,并以 Poly 方式变化^[18]。训练的最大 Epoch 设置为 300,批量大小设置为 8。另外采用 L2 正则化防止过拟合。

3 结果与分析

3.1 性能评估

由于本文涉及叶片和病斑分割以及病害严重度估算,因此采用准确率(Precision)、召回率(Recall)、交并比(Intersection over union, IoU)和 F1 值^[22]来评价分割结果。模型整体分割性能由像素精度(Pixel accuracy, PA)、平均交并比(Mean intersection over union, MIoU)和频权交并比(Frequency weighted intersection over union, FWIoU)来评估^[23]。严重度估计的准确性由决定系数(R^2)和均方根误差

(RMSE)来评价^[11]。
为了验证提出模型对黄瓜叶部病害严重度估算的准确性,分别对提出的网络进行消融实验,并且与先进的估算模型进行对比。在此期间,展示了提出模型在图像分割任务中的定性和定量结果,以及病害严重度估算的评估结果。

3.2 消融实验

为了评估 MDCB 和 CBAM 在提出的严重度估算模型中的有效性,在黄瓜病害数据集上进行了消融实验研究。使用测试集评估训练好的分割模型的性能,并通过测试分割图估算病害严重度。在模型训练中,使用原始 UNet 作为基线模型,添加 BN 层,从头开始训练。分别通过 MDCB 和 CBAM 优化 UNet,得到 UNet + MDCB 和 UNet + CBAM 2 个优化模型。最后,通过在原始 UNet 中同时配备 MDCB 和 CBAM 模块获得 MA - UNet(UNet + MDCB + CBAM)。

病害数据集上的语义分割结果如图 7 所示,田间条件采集的病害图像含有相近颜色的背景,并且病害病症有其特殊性,病斑可能存在不同尺度、不规则形状和模糊边界的现象。仅依靠原始 UNet,通常无法获得令人满意的分割结果(图 7c)。通过增加感受野的多样性和补偿传统扩张卷积中普遍存在的局部细节损失,UNet + MDCB 获得了更好的语义分割效果(图 7d)。另一方面,通过提取病害的位置信息和空间信息,克服复杂背景的影响,UNet + CBAM 也优于基线模型(图 7e)。通过在 MA - UNet 中集成 MDCB 和 CBAM 模块,不仅可以提高模型识别各种尺度和不规则形状病斑的能力,还可以增强网络对复杂背景的泛化能力(图 7f)。

除了视觉比较之外,表 1 展示了测试集上各模型的整体性能。结果显示,UNet + MDCB 和 UNet + CBAM 都获得了比原始 UNet 更优的 PA、MIoU 和 FWIoU,清楚地证明了 MDCB 和 CBAM 模块在复杂背景下分割不规则病害的有效性。基于集成了 MDCB 和 CBAM 的 UNet,本文提出的 MA - UNet 模型在测试集上以最佳的整体性能优于其他对比模型,最终 PA、MIoU 和 FWIoU 分别达到 96.79%、84.97% 和 93.95%,相比 UNet 模型,分别提高 5.08、11.78、8.97 个百分点。消融实验中各模型对病害严重度的估算情况与分割结果一致,随着 PA、MIoU 和 FWIoU 的提升,MA - UNet 在病害严重度估算任务中的拟合度升高,最终 R^2 达到 0.965 4,比原始 UNet 模型提升 0.080 3,均方根误差为 1.083 7,相比原始 UNet 降低 1.605,表明 MA - UNet 模型在估算病害严重度时偏差最小。

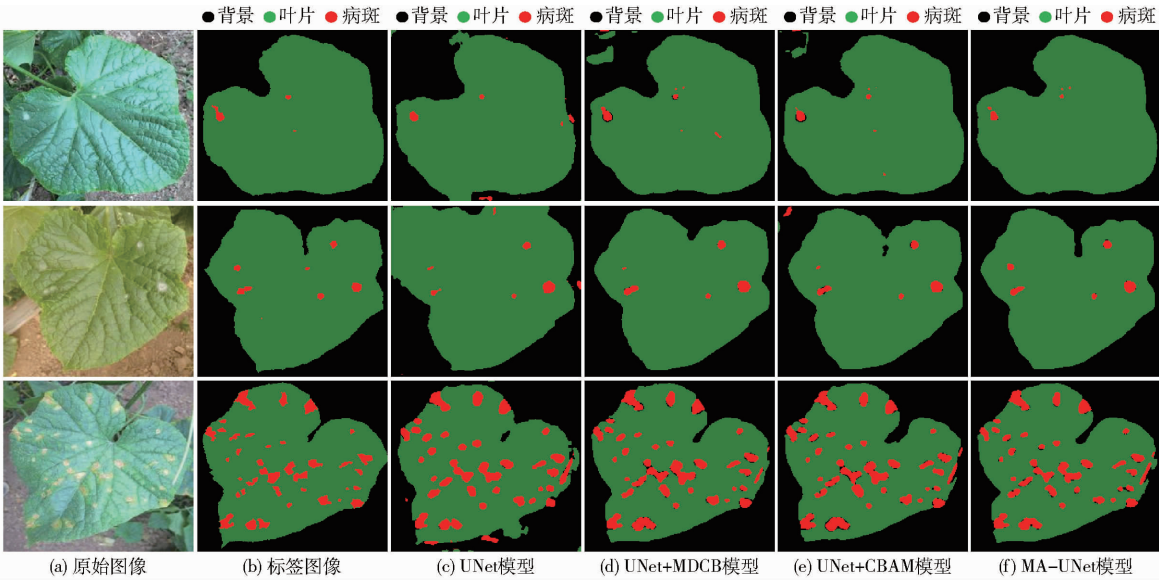


图 7 病害数据集上消融实验的视觉比较

Fig. 7 Visual comparison of ablation studies on test dataset

表 1 消融实验结果

Tab.1 Results of ablation studies

模型	PA/%	MIoU/%	FWIoU/%	R^2	RMSE
UNet	91.71	73.19	84.98	0.885 1	2.688 7
UNet + MDCB	96.58	84.58	93.57	0.963 6	1.433 3
UNet + CBAM	96.51	84.26	93.46	0.965 1	1.596 6
MA - UNet	96.79	84.97	93.95	0.965 4	1.083 7

表 2 为消融实验中各模型在背景类别、叶片类别和病斑类别上的精确率、召回率、交并比和 F1 值。结果表明,对于 3 个类别来说,消融实验中的各个模型的分割结果都优于原始 UNet。总体来看,背景类别上的分割性能均高于叶片类别和病斑类别。MDCB 模块使模型对病斑类别的 F1 值和 IoU 提升到 79.78%、66.36%,对叶片类别的 F1 值和 IoU 提升到 97.46%、95.05%,表明 MDCB 通过增大特征图感受野的同时保持病症信息的连续性和完整性,

表 2 消融实验类别结果

Tab.2 Results of ablation studies for each category

%					
模型	类别	精确率	召回率	F1 值	交并比
UNet	背景	92.46	84.65	88.38	79.19
	叶片	92.74	95.40	94.05	88.77
	病斑	59.76	79.12	68.09	51.62
UNet + MDCB	背景	96.04	95.97	96.00	92.32
	叶片	97.83	97.10	97.46	95.05
	病斑	72.11	89.27	79.78	66.36
UNet + CBAM	背景	94.76	97.21	95.97	92.25
	叶片	98.39	96.44	97.40	94.94
	病斑	71.65	88.56	79.21	65.58
MA - UNet	背景	94.58	98.12	96.32	92.90
	叶片	98.69	96.55	97.61	95.34
	病斑	75.43	85.19	80.01	66.68

能够提高像素级分类的准确性。CBAM 同样使得分割模型对 3 个类别的精确率、召回率、交并比和 F1 值都有提升,表明像素类内响应和通道间的依赖关系能够提升模型的分割效果。MA - UNet 结合了 MDCB 和 CBAM 的双重优势,进一步提升了叶片类别和病斑类别的准确率和交并比,有助于在严重度估算任务中准确地获取病斑和叶片的像素数,降低严重度估算的偏差。

3.3 模型性能对比

为了进一步验证所提出的 MA - UNet 模型对严重度估计的有效性,与最具代表性的语义分割模型进行比较,包括 FCN、SegNet、DeepLabV3 +、FPN 以及 PSPNet。为了保证比较的公平性,对比过程使用相同的训练策略和相同的计算环境。同样通过 PA、MIoU 和 FWIoU 来评价模型的分割性能,通过 R^2 和 RMSE 评价模型的严重度估算性能。

表 3 为不同语义分割模型在测试集上的性能参数对比。在分割准确率方面,本文提出的 MA - UNet 取得了比对比模型更好的性能参数,这得益于不同膨胀率的混合扩张卷积提取不同尺寸的病斑信息,CBAM 注意力使网络聚焦于叶片和病斑类别,同步提升病害分割精度。FPN 模型分割精度效果仅次于 MA - UNet,它的 PA、MIoU 和 FWIoU 分别为 94.96%、76.21% 和 90.87%。分割精度相对表现最差的是 FCN,其 PA 为 90.05%,FWIoU 仅为 82.2%。MA - UNet 的 FWIoU 为 93.95%,比表现最差的 FCN 模型高 11.75 个百分点,比 FPN 模型高 3.08 个百分点。由于估算病害严重度时仅用到叶片像素和病斑像素,所以图 8 展示了各模型在叶片类别和病斑类别上的 F1 值和交并比。对于叶片类

别,本文提出的 MA - UNet 在 F1 值和交并比上相比 FCN、SegNet、DeepLabV3 +、FPN 和 PSPNet 等对比模型至少高 4.04、5.37 个百分点。对比模型在病斑类别上的分割精度均不高,但是提出的 MA - UNet 在 F1 值和交并比上相比对比模型至少高 12.48 个百分点和 14.61 个百分点。根据表 3 显示,提出的 MA - UNet 模型在单幅图像测试中耗时 1.36 s,能够在不增加过多运算时间的情况下,取得最高的估算精度,实现精准快速的严重度估算。

表 3 不同模型性能比较

Tab.3 Performance comparison of different models						
模型	PA/ %	MIoU/ %	FWIoU/ %	R^2	RMSE/ %	单幅图像测 试时间/s
FCN	90.05	70.06	82.20	0.840 9	2.810 2	1.14
SegNet	93.09	74.59	87.44	0.831 9	2.307 6	2.46
DeepLabV3 +	92.75	74.13	86.82	0.847 7	2.736 5	1.47
FPN	94.96	76.21	90.87	0.915 4	2.160 3	1.29
PSPNet	93.81	74.87	88.89	0.905 6	2.952 3	1.35
MA - UNet	96.79	84.97	93.95	0.965 4	1.083 7	1.36

图 9 展示了所有模型的语义分割结果(为了便于视觉比较,展示与图 7 相同的测试图像)。可以看出,当 FCN 面对复杂的背景和小目标时,难以分割叶片边缘^[24]。这可能是由于 FCN 使用反卷积进行上采样,获得了稀疏的特征图,导致 FCN 对图像的细节不够敏感,丢失了像素之间的语义信息。DeepLabV3 + 采用 ASPP 模块来编码多尺度的上下文信息并抑制背景干扰^[25]。SegNet 和 DeepLabV3 + 的分割结果比 FCN 好。但是 SegNet 难以应对复杂背景的影响,叶片边界同样分割不清晰。DeepLabV3 + 采用空洞卷积提取病害多尺度特征以提高分割的准确性,但是缺少空间维度和通道维度之间的相互依赖关系,导致对病斑的分割精度略低。FPN 和 PSPNet

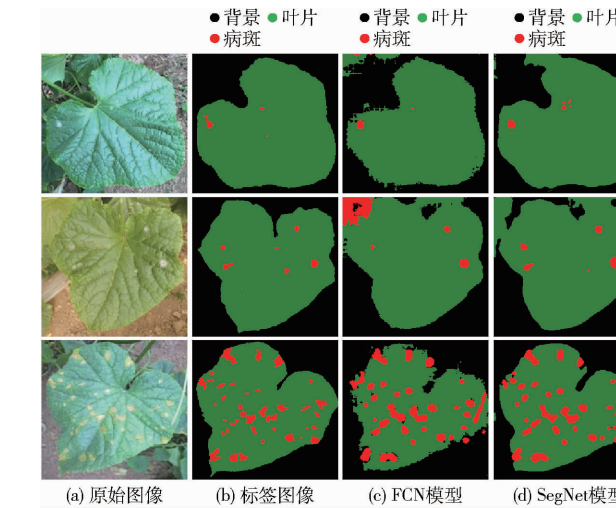


图 9 不同分割模型的分割效果

Fig.9 Segmentation results of different models on test images

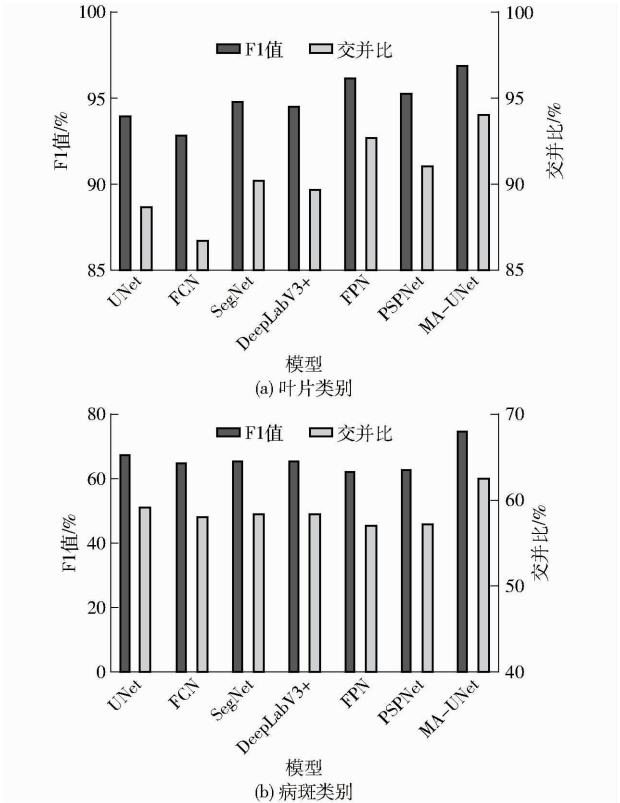
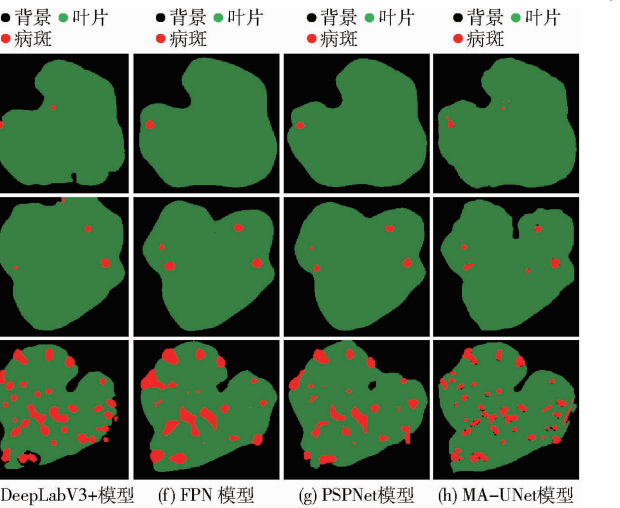


图 8 分割模型在叶片类别和病斑类别上的性能

Fig.8 Performances of segmentation models on leaf and lesion

分别通过多尺度特征图融合和聚合上下文信息对叶片精准分割,但是同样难以处理不同尺寸的病斑,导致病斑分割较差^[26]。提出的 MA - UNet使用不同膨胀率的混合扩张卷积在扩大感受野的同时能够捕捉并聚焦于不同尺寸的病斑,提取空间和通道的相互依赖关系,从而捕捉到了病害图像更细化更精准的特征^[20]。因此,MA - UNet 在叶子和病斑分割方面取得了明显改善,并有效地抵抗背景噪声的影响,减少了叶片过分割。

根据语义分割模型的分割图估算病害严重度,



对模型的可靠性进行了评估^[3,14],估计结果如图 10 所示。结果显示,所有模型的 R^2 均超过 0.83,估算精度范围从 0.831 9 (SegNet) 到 0.965 4 (MA - UNet),表明基于语义分割方法能够估计病害严重度^[3]。FCN、PSPNet 和 DeepLabV3 + 的估计误差相对较大,分别为 2.810 2%、2.952 3% 和 2.736 5%。FPN 取得的严重度估计结果稍好, R^2 为 0.915 4,但其估算的偏差较高。MA - UNet 估计严重度与实际

严重度有很好的 consistency,取得了最高的 R^2 ,为 0.965 4,最低的 RMSE,为 1.083 7%,这取决于病斑和叶片的准确分割。从图 10 总体看来,大多数模型倾向于高估病害严重度^[3,12],导致估计误差更明显。对严重度的高估可能是由于发病严重的病害叶片中病斑面积很大,还存在许多小的病斑相互粘连,导致模型对健康叶片和病斑的分割不准确,这也与人工视觉判断一致^[2]。

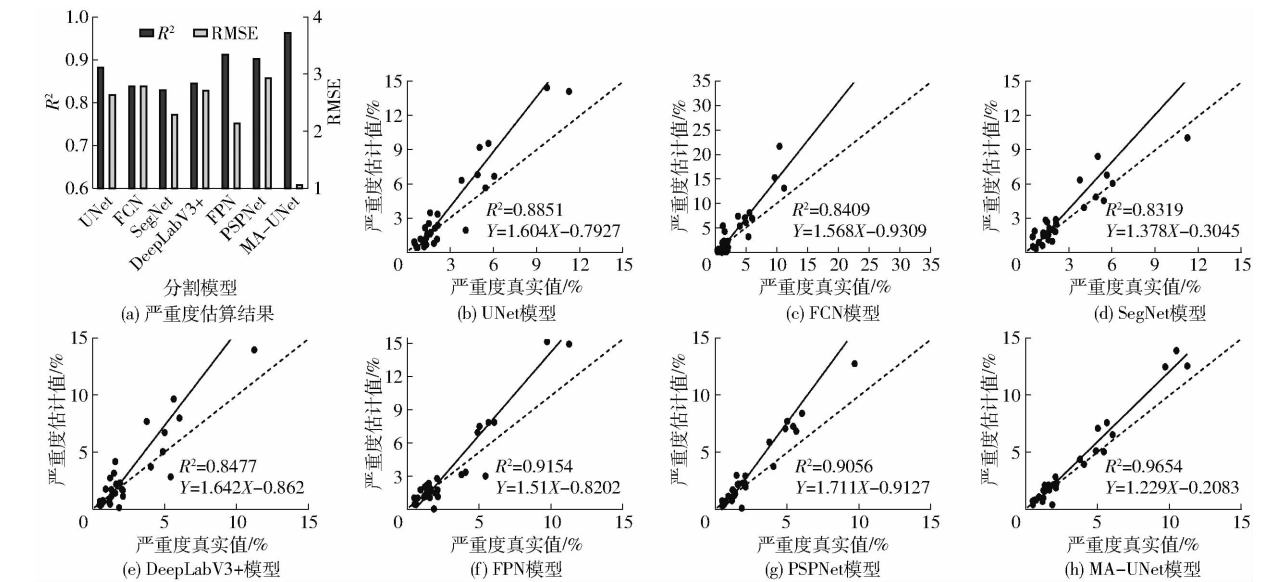


图 10 不同语义分割模型对严重度估算的结果

Fig. 10 Severity estimation results based on semantic segmentation networks

4 结论

- (1) 针对病害图像中病斑尺寸多变,设计具有不同膨胀率的混合扩张卷积块 (MDCB),通过增加感受野的多样性和最小化传统扩张卷积中网络问题,提高模型识别各种尺度和不规则形状病斑的能力。
- (2) 针对田间获取的病害图像中的复杂噪声问题,设计引入注意力机制 (CBAM),通过增强模型对关键细节信息表达,提升特征精准定位的效果,细化

病斑分割,提高严重度估计的准确性。

(3) 选择 FCN、UNet、SegNet、DeepLabV3 +、FPN 和 PSPNet 模型与本文提出的 MA - UNet 进行对比分析。在病害分割中,MA - UNet 的像素精度、平均交并比和频权交并比分别为 96.79%、84.97% 和 93.95%,均优于其他深度学习模型。基于分割结果,在病害严重度估算中,MA - UNet 模型与真实严重度的决定系数达到 0.965 4,最多比对比模型高 0.133 5, RMSE 最小,仅为 1.083 7%,表明本文模型能够对叶部病害的严重度进行精准估计。

参 考 文 献

[1] ATALLAH O O, OSMAN A, MAS A, et al. Soybean β -conglycinin and catfish cutaneous mucous p22 glycoproteins deteriorate sporangial cell walls of *Pseudoperonospora cubensis* and suppress cucumber downy mildew[J]. Pest Management Science, 2021, 77(7): 3313 - 3324.

[2] GAO J, WESTERGAARD J C, SUNDMARK E H R, et al. Automatic late blight lesion recognition and severity quantification based on field imagery of diverse potato genotypes by deep learning[J]. Knowledge-Based Systems, 2021, 214: 106723.

[3] BOCK C H, CHIANG K, DEL P E M. Plant disease severity estimated visually: a century of research, best practices, and opportunities for improving methods and practices to maximize accuracy[J]. Tropical Plant Pathology, 2022, 47(1S1): 25 - 42.

[4] XUE W, HAYNES K G, QU X. Characterization of early blight resistance in potato cultivars[J]. Plant Disease, 2019, 103(4): 629 - 637.

[5] DEL P E M, NELSON S C, PETHYBRIDGE S J. Evaluation of app-embedded disease scales for aiding visual severity estimation of *Cercospora* leaf spot of table beet[J]. Plant Disease, 2019, 103(6): 1347 - 1356.

[6] 翟肇裕,曹益飞,徐焕良,等. 农作物病虫害识别关键技术研究综述[J]. 农业机械学报, 2021, 52(7): 1 - 18.

- ZHAI Zhaoyu, CAO Yifei, XU Huanliang, et al. Review of key techniques for crop disease and pest detection[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 1 – 18. (in Chinese)
- [7] ESCARIO J G M, KROHLING R A, VENTURA J A. Deep learning for classification and severity estimation of coffee leaf biotic stress[J]. Computers and Electronics in Agriculture, 2020, 169: 105162.
- [8] LIANG Q, XIANG S, HU Y, et al. PD²SE – Net: computer-assisted plant disease diagnosis and severity estimation network[J]. Computers and Electronics in Agriculture, 2019, 157: 518 – 529.
- [9] 鲍文霞, 林泽, 胡根生, 等. 基于 RSTCNN 的小麦叶片病害严重度估计[J]. 农业机械学报, 2021, 52(12): 242 – 252. BAO Wenxia, LIN Ze, HU Gensheng, et al. Severity estimation of wheat leaf diseases based on RSTCNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(12): 242 – 252. (in Chinese)
- [10] 张领先, 田潇, 李云霞, 等. 可见光光谱和机器学习的温室黄瓜霜霉病严重度定量估算[J]. 光谱学与光谱分析, 2020, 40(1): 227 – 232. ZHANG Lingxian, TIAN Xiao, LI Yunxia, et al. Estimation of disease severity for downy mildew of greenhouse cucumber based on visible spectral and machine learning[J]. Spectroscopy and Spectral Analysis, 2020, 40(1): 227 – 232. (in Chinese)
- [11] RONNEBERGER O, FISCHER P, BROX T. U – Net: convolutional networks for biomedical image segmentation[J]. Medical Image Computing and Computer-Assisted Intervention, PT III, 2015, 9351: 234 – 241.
- [12] WSPANIALY P, MOUSSA M. A detection and severity estimation system for generic diseases of tomato greenhouse plants[J]. Computers and Electronics in Agriculture, 2020, 178: 105701.
- [13] GONCALVES J P, PINTO F A C, QUEIROZ D M, et al. Deep learning architectures for semantic segmentation and automatic estimation of severity of foliar symptoms caused by diseases or pests[J]. Biosystems Engineering, 2021, 210: 129 – 142.
- [14] WANG C, DU P, WU H, et al. A cucumber leaf disease severity classification method based on the fusion of DeepLabV3 + and U – Net[J]. Computers and Electronics in Agriculture, 2021, 189: 106373.
- [15] 张垵琦, 亢宇鑫, 武卓越, 等. 基于多尺度特征和注意力机制的肝脏组织病理图像语义分割网络[J]. 模式识别与人工智能, 2021, 34(4): 375 – 384. ZHANG Aoqi, KANG Yuxin, WU Zhuoyue, et al. Semantic segmentation network of pathological images of liver tissue based on multi-scale feature and attention mechanism[J]. Pattern Recognition and Artificial Intelligence, 2021, 34(4): 375 – 384. (in Chinese)
- [16] 饶秀勤, 朱逸航, 张延宁, 等. 基于语义分割的作物垄间导航路径识别[J]. 农业工程学报, 2021, 37(20): 179 – 186. RAO Xiuqin, ZHU Yihang, ZHANG Yanning, et al. Navigation path recognition between crop ridges based on semantic segmentation[J]. Transactions of the CSAE, 2021, 37(20): 179 – 186. (in Chinese)
- [17] 马保建, 鄢金山, 王乐, 等. 基于语义分割的矮化密植枣树修剪枝识别与骨架提取[J]. 农业机械学报, 2022, 53(8): 313 – 319, 442. MA Baojian, YAN Jinshan, WANG Le, et al. Method for detection and skeleton of pruning branch of jujube tree based on semantic segmentation for dormant pruning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8): 313 – 319, 442. (in Chinese)
- [18] CHEN L, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[J]. Computer Vision-ECCV 2018, PT VII, 2018, 11211: 833 – 851.
- [19] JIANG W, LIU M, PENG Y, et al. HDCB – Net: a neural network with the hybrid dilated convolution for pixel-level crack detection on concrete bridges[J]. IEEE Transactionson Industrial Informatics, 2021, 17(8): 5485 – 5494.
- [20] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module[J]. Computer Vision-ECCV 2018, PT VII, 2018, 11211: 3 – 19.
- [21] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactionson Pattern Analysisand Machine Intelligence, 2017, 39(12): 2481 – 2495.
- [22] MA J, LI Y, LIU H, et al. Improving segmentation accuracy for ears of winter wheat at flowering stage by semantic segmentation[J]. Computers and Electronics in Agriculture, 2020, 176: 105662.
- [23] 钟昌源, 胡泽林, 李森, 等. 基于分组注意力模块的实时农作物病害叶片语义分割模型[J]. 农业工程学报, 2021, 37(4): 208 – 215. ZHONG Changyuan, HU Zelin, LI Miao, et al. Real-time semantic segmentation model for crop disease leaves using group attention module[J]. Transactions of the CSAE, 2021, 37(4): 208 – 215. (in Chinese)
- [24] XU L, LI Y, XU J, et al. Two-level attention and score consistency network for plant segmentation[J]. Computers and Electronics in Agriculture, 2020, 170: 105281.
- [25] ZHANG D, DING Y, CHEN P, et al. Automatic extraction of wheat lodging area based on transfer learning method and DeepLabv3 + network[J]. Computers and Electronics in Agriculture, 2020, 179: 105845.
- [26] NGUGI L C, ABDELWAHAB M, ABO-ZAHHAD M. Tomato leaf segmentation algorithms for mobile phone applications using deep learning[J]. Computers and Electronics in Agriculture, 2020, 178: 105788.