Research on Semantic Segmentation of Crop Diseases Based on Improved Unet Model*

LIU Weili, SUN Gonglingyun, LIAN Junbo, LI Xiangchen, HUI Guohua*
(School of Mathematics and Computer Science, Zhejiang Agriculture and Forestry University, Hangzhou Zhejiang 31300, China)

Abstract: Targeting at the challenges posed by high intra-class variability and significant inter-class similarity in crop diseases, an enhanced UNet model called VGSA-Unet is presented to improve the accuracy of segmentation for crop leaf disease images. VGG16 is employed as the encoder for the Unet model. A multi-scale context module of GC_ASPP is integrated between the encoder and decoder. This addition enhances the extraction of global features from images while capturing multi-scale information, enabling the model to comprehend the details of disease images with varying morphologies and improve performance when processing images with significant scale variations. In the decoder, bilinear interpolation is used in place of transposed convolution, reducing the loss of edge information and boosting the model's computational efficiency. A hybrid loss function is applied to enhance the model's capability to segment minute disease spots. Experimental results demonstrate that the improved VGSA-Unet model outperforms existing network models in terms of mPA (92.04%) and mIoU (84.61%) metrics. The findings confirm that the proposed model overcomes the segmentation difficulties caused by disease characteristics, accurately extracts small disease spots and effectively performs precise segmentation of crop leaf diseases.

Key words: semantic segmentation; crop diseases; deep learning; Unet; GC block; atrous spatial pyramid pooling
EEACC: 7230 **doi**: 10.3969/j.issn.1004-1699.2025.10.012

基于改进 Unet 模型的农作物病害语义分割研究*

刘伟立,孙公凌云,连俊博,李响陈,惠国华*
(浙江农林大学数学与计算机科学学院,浙江 杭州 11300)

摘 要:针对农作物病害类内差异性大、类间相似性高等特点导致的分割困难问题,提出了一种改进的 Unet 模型 VGSA-Unet,用于提高农作物叶片病害图像的分割准确性。将 VGG16 作为 Unet 模型的编码器,在编码器与解码器之间集成了多尺度上下文模块 GC_ASPP,在获取图像多尺度信息的同时提升对图像全局特征的提取能力,使模型理解不同形态的病害图像细节,提高模型在处理尺度变化较大的图像时的性能;在解码器中,使用双线性插值代替转置卷积,减少边缘信息的损失,提高模型计算效率;使用混合损失函数,提高模型对细小病斑的分割能力。与其他网络模型相比,改进的 VGSA-Unet 模型在 mPA (92.04%) 和 mIoU (84.61%) 两项指标上均达到了最高值。改进后的模型克服了病害特点造成的分割困难,对细小病斑的提取更为精确,能有效地对农作物叶片病害进行准确分割。

关键词:语义分割;农作物病害;深度学习;Unet;GC block;空洞空间金字塔池化
中图分类号:TP391 **文献标识码**:A **文章编号**:1004-1699(2025)10-1818-09

农作物病害是世界各地农业生产中的主要问题之一。病害对农作物的生长、发育和产量造成了严重的影响[1-2]。农作物叶片病害图像分割可以帮助及时准确地识别和监控病害,从而采取有效措施防治病害、提高农作物产量和质量。然而,这一任务面临诸多挑战,包括病害形态多样性、类内差异性大、类间相似性高等问题[3],这些因素都给准确分割病害区域带来困难。

卷积神经网络具有局部感知性、参数共享、逐层提取特征以及自动学习表示等特点,广泛应用于图像识别及语义分割等领域。在图像分类领域,AlexNet[4]、VGGNet[5]、GoogLeNet[6] 和 ResNet[7] 等卷积神经网络模型都取得较好的效果。最新研究中,图像分类方法已在农作物病害识别领域取得瞩目成就。李好等[8]提出一种改进 ShuffleNet V2 的农作物病害识别方法,在 PlantVillage 数据集上的分类准确率达到了 99.24%。在当下研究成果基础上,继续提高识别准确率较为困难,因此研究重心已经转到模型轻量化方向。王焕鑫等[9]为了能在移动端实现叶片病害识别,改进了 MobileNetV2 模型,通过剪

项目来源:国家自然科学家基金项目(U1709212)
收稿日期:2023-09-14 **修改日期**:2023-12-14

枝操作进行模型轻量化处理,最终在模型参数量减少的同时使准确率提升到 98.4%。病害分类的准确率已达到了较高水准,但仍存在着一定的局限性,例如一片叶子中出现多种不同种类的病害等类似问题,仅靠分类方法无法解决,且在实际应用中,往往达不到理论上的准确率,给农业生产带来不便。相比于目前主流的图像分类方法,语义分割任务允许在一张图像中同时识别和分割出多个物体,其可以将图像中的每个像素分配到特定的语义类别的特点,对于分割出物体的轮廓和边界起到关键作用。因此,对病害部位进行语义分割至关重要。

在语义分割领域,SegNet[10]、Unet[11]、pspnet[12]、Deeplabv3+[13-14]等模型也取得了不俗的成绩。在现有农作物病害识别研究中,关于语义分割的文献报道相对较少。钟昌源等[15]提出一种基于分组注意力模块的农作物病害叶片语义分割模型,验证了注意力模块对农作物病害叶片语义分割的有效性。但是由于其只是分割出生病的叶片,不能充分体现模型的优势,限制了模型的应用范围。赵小虎等[16]在 Unet 网络结构基础上进行改进,加入注意力机制模块,实现多尺度提取番茄病害特征。研究结果表明,多尺度提取特征能有效地提高番茄病害分割性能,但是当面对不同作物的不同病害分割时,不同作物的不同病害类型可能表现为相似的症状,这时只依靠提取多尺度特征无法解决问题。相比于文献[15]中分割整个病害叶片和文献[16]中分割单一种类作物,本研究对多种作物病害类型和病斑部位进行分割,能解决前者存在的问题,分割结果细节更丰富,具有潜在的优势。

本研究基于改进 Unet 模型,提出了用于农作物病害叶片语义分割的 VGSA-Unet 模型。该模型将 VGG16 作为特征提取器,设计了融合全局上下文模块和多尺度特征模块的 GC_ASPP 模块,以更好地捕捉和处理复杂病害特征。此外,模型采用混合损失函数,改善了边界细节缺失的问题。基于以上改进,VGSA-Unet 不但可以提高农作物叶片病害图像的分割准确性,也可以有效应对病害类内差异性大、类间相似性高导致的分割准确率低的问题。

1　材料与方法

1.1　数据集的建立

1.1.1　数据采集及标注

PlantVillage 数据集[17]是一个旨在促进植物病理学和自动化诊断系统研究的大型图像数据集。该数据集由美国宾夕法尼亚州立大学的研究人员开发,包含来自 14 种不同植物的图像。本研究以苹果和葡萄两种农作物为研究对象,包括苹果黑星病、褐斑病、炭疽病以及葡萄黑腐病、褐斑病、轮斑病等共 6 类病害,如图 1 所示。本文使用的数据集从 PlantVillage 数据集对应分类中选取每类病害样本各 300 张,共 1 800 张高质量图片。使用图像标注工具 Labelme 对收集的病害图像病害区域进行人工手动标注,使用多种标注工具如多边形、矩形或笔刷,在图像上绘制区域并为其分配语义标签。完成标注后,保存标注结果为 JSON 文件,其中包含了图像路径、标注的区域和对应的语义标签,最后将构建好的数据集命名为“PV6”。

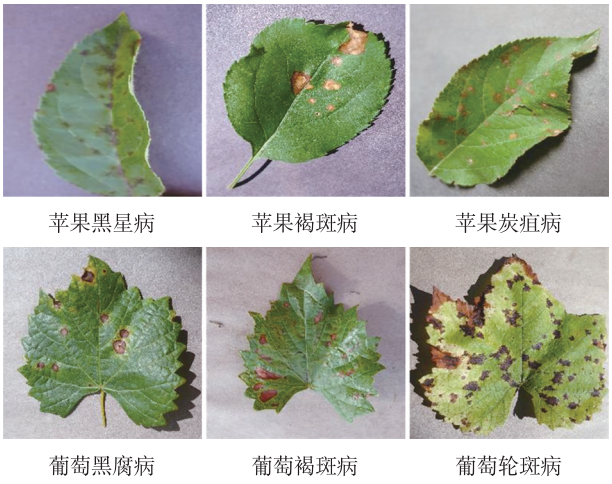

图 1　PlantVillage 图片集示例图

1.1.2　数据特点

对于农作物病害而言,叶片病害表现形式多种多样,包括斑点、斑块、褐化、变形等。此外,农作物的叶片病害也具有其特殊性,例如:①类内差异性大:同一种病害在不同的阶段或在不同的环境条件下可能呈现不同的表现形态。病害开始时可能只是一些小斑点,但随着病害的发展,这些斑点可能会扩大、合并,或者颜色加深。②类间相似性高:不同种类的病害可能表现相似症状,例如都出现黄斑或者褐斑。这使得仅依靠局部特征很难区分不同种类的病害。③病害区域大小和位置的多样性:病害可能在叶片的任何位置出现,大小也可能从几个像素到几百个像素不等。

1.1.3　数据预处理

将数据按照 9:1 的比例划分训练集和验证集。在训练模型之前,为了提高模型的训练效果和鲁棒性,使其能够更好地适应不同场景和数据的特点,对图像进行预处理。首先将输入图像的分辨率统一调整为 512×512 像素,然后对每个批次的图像进行线上图像增强,包括图像长宽比扭曲、图像翻转、图像

裁剪、随机缩放和色域变化方法,增强了数据的多样性。其中通过色域变化方法对图像进行处理,可以模拟不同的光照条件和颜色偏差,提高了模型对于实际应用场景中光照和颜色变化的适应性。最后对图像样本统一进行归一化处理。图 2 展示了(以图 1图示为例)经过增强处理的某个批次图像。

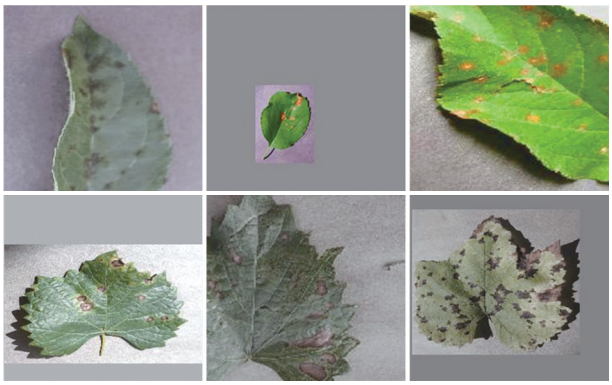

图 2 图像增强示例图

1.2 模型的改进

1.2.1 改进的 Unet 模型

针对农作物叶片病害的特点,对 Unet 网络进行了改进,将改进后的模型命名为 VGSA-Unet,其结构如图 3 所示。对比普通卷积、残差网络、深度可分离卷积的结果后,选择效果最佳的 VGG16 作为主干网络,删除其最后的全连接层,仅保留特征提取部分作为编码器。输入图像可以是任意大小的农作物病害图像,在输入网络之前,首先对图像进行预处理操作。

编码器部分位于图 3 的左侧。该部分进行了 4 次 2 倍下采样,每次下采样都将输入特征图的尺寸大小缩减为原来的一半。图像输入网络,经过主干网络对原始图像进行特征提取,得到四种不同尺度的特征图(1/2,1/4,1/8,1/16),然后将最深层特征送入 GC_ASPP 模块。

解码器部分位于图 3 的右侧。该部分采用了自定义的上采样模块,每个模块包括上采样层、卷积层和激活函数。在上采样阶段,使用双线性插值来替代转置卷积,以减小边缘信息的损失并提高计算效率。解码阶段从 GC_ASPP 模块获得的特征图开始,逐步进行上采样和特征融合。在解码器的最后,通过一个 1×1 的卷积层得到最终的输出。

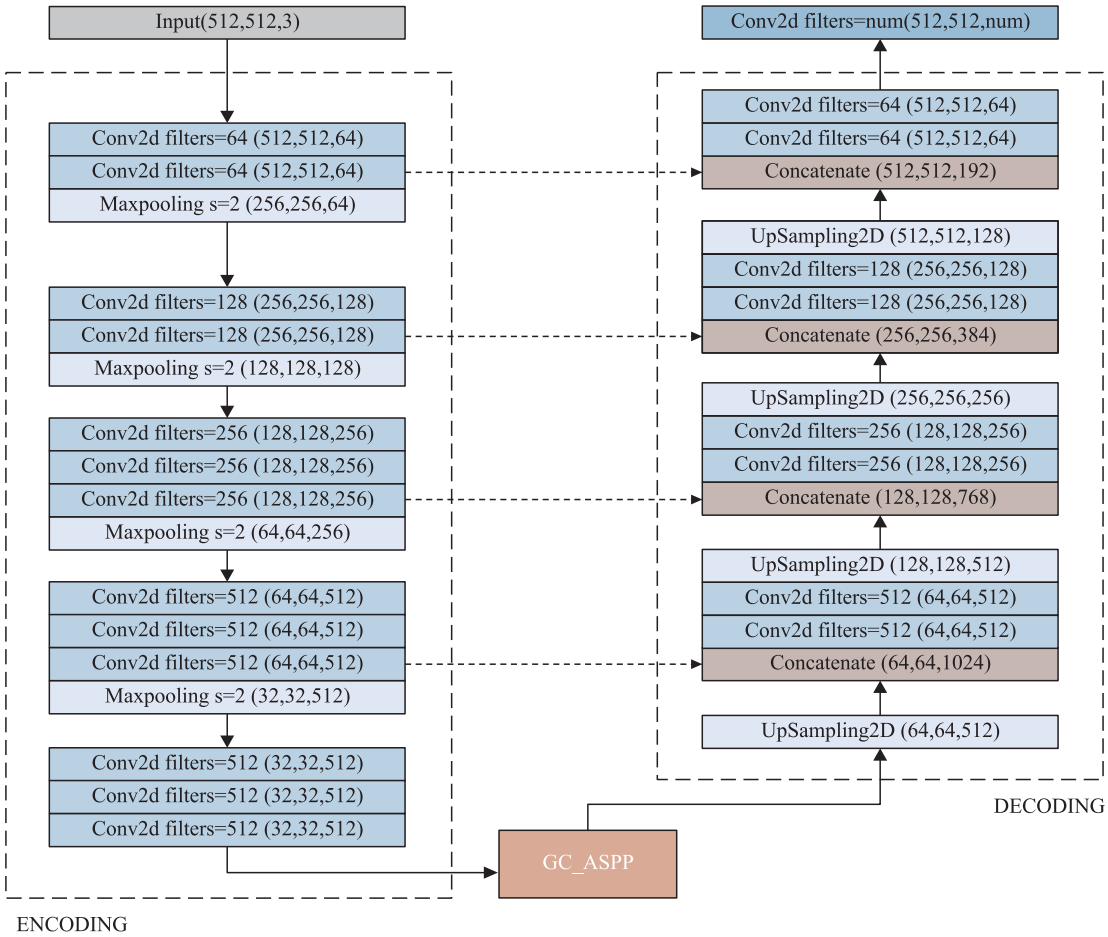

图 3 VGSA-Unet 模型结构图

1.2.2　空洞卷积与空间金字塔池化

在普通卷积层中,每个输出神经元只与输入图像中某个小区域内的神经元相连,这个区域被称为卷积核的感受野。在深度卷积神经网络中,随着网络层数的加深,感受野逐渐变大,但增长速度非常缓慢。因此,对于大尺寸的输入图像来说,最后一层的感受野较小,无法很好地捕捉全局信息。空洞卷积[18]通过在卷积核中引入空洞来增加感受野的大小,从而解决了上述问题。

空洞空间金字塔池化(ASPP[19]模块)的核心思想是利用空洞卷积在不同的采样率上同时进行卷积操作,以捕捉和融合不同尺度的信息。相较于单一尺度的空洞卷积,ASPP 可以同时利用多个不同采样率的空洞卷积来捕捉不同尺度下的特征信息。这使得 ASPP 具有更强的多尺度特征提取能力,更好地捕捉物体在不同尺度上的形状和细节信息;ASPP 采用金字塔结构的特征池化方法,在不同的空间采样率下对特征图进行池化,从而在保留空间信息的同时,更好地融合全局的语义信息;ASPP 通过在不同的空洞率下进行卷积操作,有效地增加感受野的大小,提高网络的视野,同时不会显著增加计算复杂度。如图 4 所示,ASPP 包含了 4 个空洞卷积层,其中有 1 个卷积核大小为 1×1 的空洞卷积层,用于捕获图像的局部特征;有 3 个卷积核大小为 3×3 的空洞卷积层(采样率分别为 6、12、18),用于捕获更大范围的感受野;最右侧包含 1 个自适应池化层,将特征图压缩至 1×1,然后通过双线性插值恢复分辨率(最右侧 3 层)。最后将所有空洞卷积分支和全局池化分支的特征在通道维度上进行拼接,经过 1 个卷积核大小为 1×1 的卷积层进行降维,得到最终特征。而池化金字塔各层的膨胀因子可自定义,从而实现自由的多尺度特征提取。

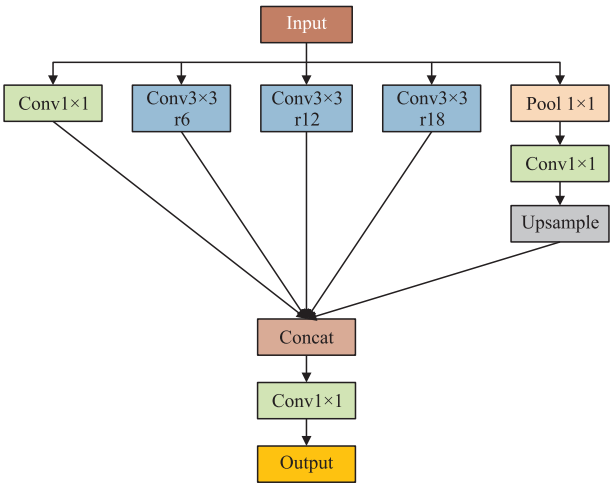

图 4　ASPP 结构图

1.2.3　全局上下文模块

在传统的 CNN 中,卷积操作只关注输入特征图中的局部区域,无法很好地捕捉全局上下文信息,可能会限制网络在复杂场景进行分类、检测和分割等任务的性能。GCNet[20]在卷积神经网络中加入全局上下文模块(GC block)(图 5),GC block 主要由语义模块(Context Modeling)和转换模块(Transform)组成。其中,语义模块首先使用卷积核大小为 1×1 的卷积层将输入特征图通道数转换为 1,表示输入特征的全局信息,然后经过 Softmax 激活函数在通道维度上对全局信息进行归一化。转换模块中,特征图通过 1 个卷积核大小为 1×1 卷积层,将输入特征图的通道数转换为全局上下文特征的通道数;接着经过 LayerNorm 层,对特征图在通道维度上进行归一化;再经过 ReLU 激活函数,引入非线性特性;然后通过 1 个卷积核大小为 1×1 的卷积层,将全局上下文特征的通道数转换回到原始输入特征的通道数。最后,将全局上下文特征与原始输入特征相加,得到最终的输出。这样,GCNet 可以有效地利用全局信息来指导网络的特征提取和分割等任务,从而提高网络的性能。

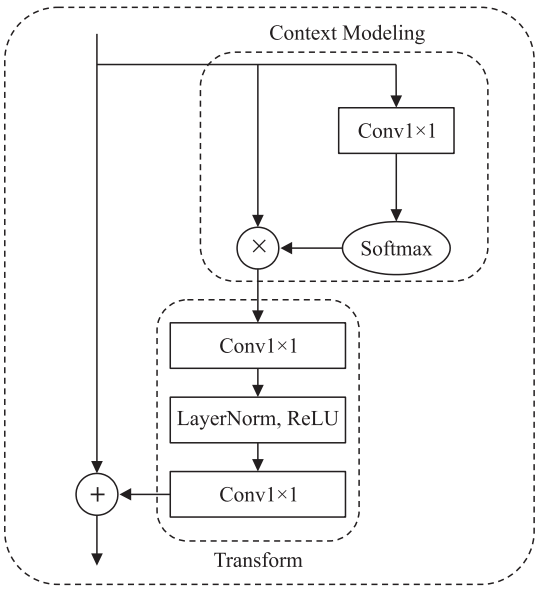

图 5　GC block 结构图

1.2.4　改进的 GC_ASPP 结构

本研究结合全局上下文模块(GC block)和空洞空间金字塔池化(ASPP)设计了 GC_ASPP 模块(图 6),以提升模型在农作物病害图像分割中的性能。GC_ASPP 模块结构通过使用不同采样率的空洞卷积层来捕获多尺度空间以及全局上下文信息,来加强模型对全局信息的理解并增强其处理不同尺度病害特征的能力。这种结合使模型不仅能理解病害的扩散规律和叶片的整体健康状态,而且在处理

具有高类内差异性和类间相似性的病害图像时，能有效预测并提高对病害区域的检测和分割能力。

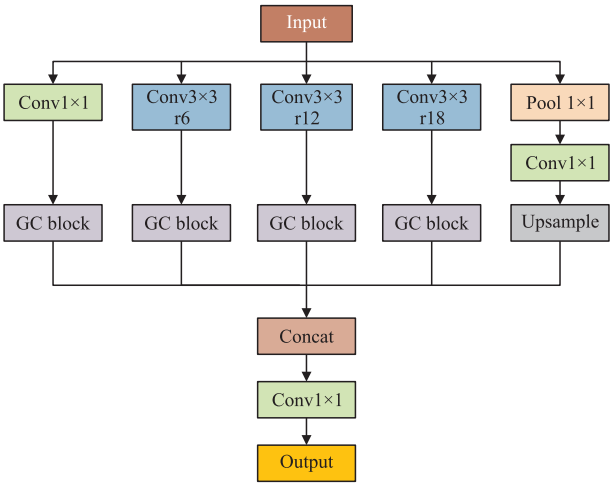

图 6 GS_ASPP 结构图

为了避免全局上下文信息在多尺度特征捕获中的不足，使每个尺度的特征都能充分整合全局上下文信息，本研究在 ASPP 中的每个空洞卷积层后面都加入了一个 GC block 模块。在 GC_ASPP 模块中，特征信息经过带有 GC block 的四个空洞卷积层处理后，得到具有不同空间分辨率的四个特征图，然后将这四个特征图与经过全局池化层处理后的全局特征图在通道维度上进行拼接，得到最终的特征图。这有助于模型更好地理解目标区域与背景区域之间的关系，从而提高分割质量；使得不同尺度特征都有机会与全局上下文信息进行融合，帮助模型更好地捕捉具有不同大小和形状的目标区域，并提高分割精度。

1.2.5 复合损失函数

CE Loss[21]（交叉熵函数）是一种经典的损失函数，广泛应用于各种分类任务，在语义分割中也取得了良好的效果。CE Loss 更关注像素级别的分类，无法直接优化边界和细节信息的预测结果，因此在处理分割任务中可能存在一定的局限性。Dice Loss[22]是一种用于语义分割任务的损失函数，通过比较预测结果与真实标签之间的重叠度来度量模型的性能。其基于 Dice 系数（也称为 Sørensen-Dice 系数），该系数衡量了两个集合的相似度。通过最小化 Dice Loss，模型被鼓励更好地学习到分割结果的边界和细节信息，从而提高语义分割模型的性能和分割质量。Dice Loss 的具体形式见式（1），CE Loss 的具体形式见式（2）。本文设计的损失函数通过结合 Dice Loss 和 CE Loss，以解决分类精度和边界细节缺失的问题，从而提高语义分割模型的性能和分割质量。

Dice Loss 的损失率 L_D 为：

$$L_D(p,q)=1-\frac{2\times\sum_{i=1}^{W}\sum_{j=1}^{H}|p_{ij}\cap q_{ij}|}{\sum_{i=1}^{W}\sum_{j=1}^{H}(|p_{ij}|+|q_{ij}|)} \tag{1}$$

CE Loss 的损失率 L_{CE} 为：

$$L_{CE}(p,q)=-\log_2\left[\frac{\exp(q_{ij})}{\sum_{i=1}^{W}\sum_{j=1}^{H}\exp(p_{ij})}\right] \tag{2}$$

混合损失函数的损失率 L_{DCE} 为：

$$L_{DCE}=L_{CE}(p,q)+L_D(p,q) \tag{3}$$

式中：W、H 分别代表输入图片的高度和宽度，p_{ij} 为模型预测结果中坐标 (i,j) 的输出值，q_{ij} 为标签中坐标为 (i,j) 真实值。

1.3 试验工具与环境

1.3.1 试验环境

本研究在 64 位的 Windows 10 操作系统上进行试验，搭载的处理器为 12th Gen Intel(R) Core(TM) i5-12600KF 3.70 GHz，显卡型号为 RTX 3080，使用 Pythorch 1.10.2 深度学习框架在 Pycharm 2021.1.2 开发环境中进行试验，使用的编程语言为 Python 3.6.5。

1.3.2 参数设置

经过多次实验调参，选择最优的学习率、优化器、Epochs、Batch Size 参数。设置初始学习率为 0.0001，并使用 Adam[23]作为优化器。采用余弦退火方法调整学习率，有助于加快模型的收敛速度并避免陷入局部最优解。训练过程中的 Epochs 设置为 150，受试验环境限制，Batch Size 设置为 8。

1.3.3 评估标准

本研究针对农作物数据集进行分割性能评估，选取平均像素准确率（mPA）、平均交并比（mIoU）、精确率（Precision）作为评价网络分割性能的指标。

mPA 计算了所有像素中正确分类的像素的比例，是每个类别的像素精度的平均值，计算公式如下：

$$\mathrm{mPA}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \tag{4}$$

mIoU 是一种常用于评估图像分割任务的度量标准，一个类别的预测区域（与实际区域重叠的部分）和实际区域的并集的比例，再求所有类别交并比的平均值，即平均交并比，计算公式如下：

$$\mathrm{mIoU}=\frac{\mathrm{TP}}{\mathrm{FP}+\mathrm{FN}+\mathrm{TP}} \tag{5}$$

Precision 计算的是所有被模型预测为正例的样本中，真正为正例的样本的比例，计算公式如下：

$$\mathrm{Precision}=\frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} \tag{6}$$

式中:TP(True Positives)表示被正确地分类为正例的像素数量;TN(True Negatives)表示被正确地分类为负例的像素数量;FP(False Positives)表示被错误地分类为正例的像素数量;FN(False Negatives)表示被错误地分类为负例的像素数量。

2　结果与分析

2.1　不同改进方式的效果比较与分析

为了验证各个模块的作用,在采用 VGG16 作为主干网络的 Unet 模型的基础上分别添加不同模块进行对比。设计方案 1 为不加入任何模块的试验;方案 2 为只加入 ASPP 模块的试验;方案 3 为只加入 GC block 的试验;方案 4 为将 GC block 模块加入 ASPP 全连接层;方案 5 为在 ASPP 中的每一个空洞卷积层后面加入一个 GC block 模块。表 1 展示了 5 种方案在 PV6 数据集上的试验结果。

表 1　各模块消融试验　　%

方案	mPA	mIoU	Precision
1	88.47	82.23	91.82
2	91.72	84.11	90.80
3	91.40	83.86	90.82
4	91.64	83.61	90.27
5	92.04	84.61	91.08

方案 2 的 mIoU 值为 84.11%,说明 ASPP 模块在提升模型性能方面具有一定的作用。ASPP 模块通过使用不同尺度的空洞卷积层能够捕捉多尺度的农作物叶片信息,从而提高分割性能。方案 3 的 mIoU 值为 83.86%,这表明 GC block 对于模型性能的提升有一定贡献。GC block 通过全局上下文信息来增强局部特征,可以帮助模型理解病斑与叶片的联系。方案 4 的 mIoU 值为 83.61%,此结果低于单独使用 ASPP 模块和 GC block 的结果。这可能是由于将 GC block 加入到 ASPP 全连接层后,使得全局上下文信息在 ASPP 模块中的整合不够充分,导致性能下降。方案 5 的 mIoU 值为 84.61%,结果优于其他方案。这表明在每个空洞卷积层后添加 GC block 模块能够更好地整合全局上下文信息,从而使 ASPP 模块在捕获多尺度上下文信息方面更加高效。

根据消融试验结果,通过对比 GC_ASPP 模块加入前后的模型性能,模型在平均像素准确率(mPA)、平均交并比(mIoU)等关键指标上均有所提升,表现出了更高的准确性和鲁棒性。这说明 GC_ASPP 模块通过结合多尺度空洞卷积和全局上下文信息可以提升模型处理不同形态病害的能力,显示了 GC_ASPP 模块在处理具有复杂特征的病害图像中的有效性。此外,消融试验进一步展示了 GC block 和 ASPP 在提升模型理解图像全局信息和细节特征方面的互补作用。

2.2　不同模型的效果比较与分析

为了验证网络结构的有效性,分别使用相同参数在 PV6 数据集上训练 FCN[24]、Deeplabv3+、pspnet、Unet、Orcnet[25]、Segformer[26]、Knet[27] 和 VGSA-Unet。VGSA-Unet 与其他常用模型处理的对比试验结果见表 2。由表 2 可知,VGSA-Unet 的平均像素准确率 mPA 值达到 92.04%,平均交并比 mIoU 值达到 84.61%,精确率 Precision 值达到 91.08%。与上述模型相比,改进的 VGSA-Unet 像素准确率提高了 4.38%~9.66%,平均交并比提高了 3.50%~8.17%。其中与未改进的 Unet 相比,VGSA-Unet 在像素准确率、平均交并比和精确率上分别提高了 5.73%、7.63%、3.92%,证明了 VGSA-Unet 在农作物病害分割问题上的优越性。

表 2　不同模型对比试验结果　　%

网络模型	mPA	mIoU	Precision
FCN	82.38	77.49	90.94
Deeplabv3+	87.46	77.50	86.61
pspnet	86.81	76.44	85.93
Unet	86.31	76.98	87.16
Orcnet	86.06	78.89	90.43
Segformer	87.30	80.52	90.94
Knet	87.66	81.11	91.35
VGSA-Unet	92.04	84.61	91.08

改进后的模型(VGSA-Unet)在训练过程中显示出的训练集和验证集的 Loss 曲线如图 7 所示,可以看出两条曲线几乎重合。这表明模型在训练数据上没有出现过拟合,并具有良好的泛化能力。

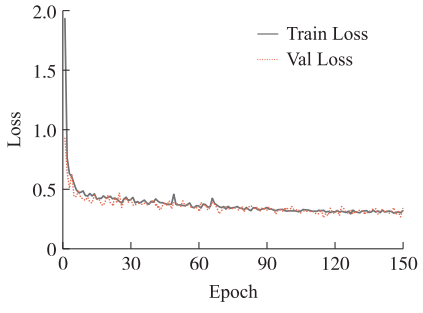

图 7　改进后模型 Loss 曲线

为更加直观地验证本方法在农作物叶片病害分割中的优势,对叶片病害分割结果进行可视化,如图 8 所示。农作物病害图像包含了许多细小病斑,因此图 9 放大了图像细节以实现更清晰的可视化。从图中可以看出,苹果褐斑病和葡萄轮斑病表现出较强的类内差异性,即病害部位既能呈现为规则斑点,又能呈现为连续不规则区域。在分割大面积连

续不规则区域时(如图 9 左侧方框),可以明显看到除了本文模型外其余模型均不能将该区域分割完整,并且不能有效地分割出相邻病斑间的轮廓。葡萄黑腐病与葡萄褐斑病表现出较强的类间相似性,即两种类型病害表现出外观相似的病害形态,改进后的网络模型能更加精准地区分相应的病害类型,对图像病害区域的分割边界更平滑,在细小病斑区域分割效果更好。

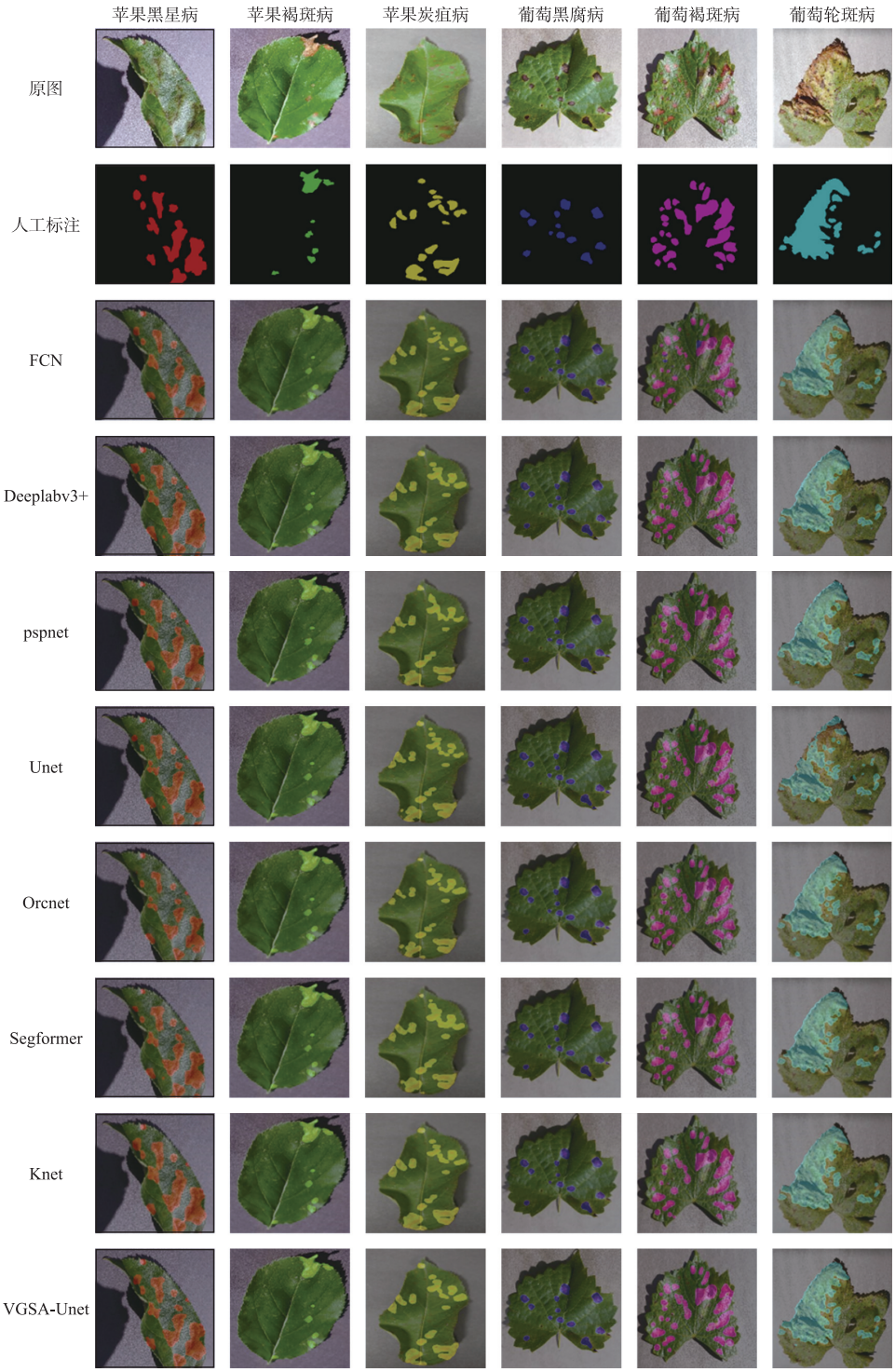

图 8 农作物病害图像分割结果

2.3 模型泛化能力验证

由于缺乏专门针对农作物病害语义分割的公开数据集,因此有必要验证所提出模型的泛化能力。为实现这一目标,使用相同参数在著名的 Pascal VOC 2007 数据集[28]上进行了试验。Pascal VOC 2007 数据集提供了 20 种物体类别和丰富的图像内容,可以在没有专用农业病害分割数据集的情况下为评估模型的性能和泛化能力提供合理的基准。最终得到的指标如表 3 所示。

由表 3 可知,改进模型 VGSA-Unet 的结果整体

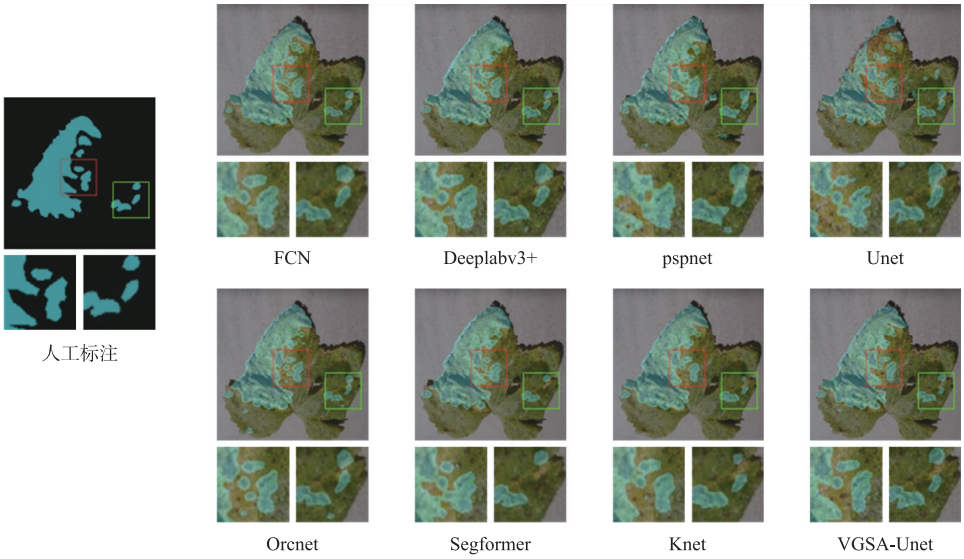

图 9　农作物病害图像分割结果(大面积不连续区域)

优于试验中的其他分割模型,虽然 mPA 略低于部分模型,但是仍处于较高水平,表明 VGSA-Unet 在像素级别的分类准确性方面具有较强的能力;mIoU 与 Precision 达到了 66.67%、81.80%,均为最高,这说明 VGSA-Unet 在区分不同类别物体的边界、正确识别目标物体等方面具有更好的性能。

以上结果表明,在 Pascal VOC 2007 数据集上,VGSA-Unet 的泛化能力较强,具有较高的语义分割性能。这主要归功于模型结构的优势,这些组件共同提高了模型在特征提取、多尺度特征融合和全局上下文信息等细节捕捉方面的能力,从而使 VGSA-Unet 在泛化能力上表现出色。

表 3　VOC 数据集对比试验结果　%

网络模型	mPA	mIoU	Precision
FCN	66.80	56.30	78.61
Deeplabv3+	78.13	63.35	74.84
pspnet	72.40	59.12	73.37
Unet	68.43	58.17	76.51
Orcnet	77.37	61.69	73.85
Segformer	77.67	61.67	73.92
Knet	71.28	61.33	78.89
VGSA-Unet	76.30	66.67	81.80

3　结论

本研究提出了一种基于改进 Unet 的农作物病害分割方法,旨在提高农作物病害图像分割的准确率。本方法采用 VGG16 作为特征提取网络,引入改进的 GC_ASPP 模块,获取图像多尺度上下文信息。设计优化的损失函数能够较好地分割出细小病斑的轮廓。改进后的模型 mPA 值达到了 92.04%,mIoU 值达到了 84.61%,有效地克服了农作物病害类内差异性大、类间相似性高所带来的分割困难等问题,表现出较高的分割性能,验证了基于改进 Unet 模型的农作物病害语义分割研究方法具有较好的应用价值。

本研究方法仍存在一定的局限性。由于所选取的数据集来自于 PlantVillage,数据集质量较高,且拍摄条件光线良好,叶片背景较为单一,因此与实地拍摄的图片在复杂背景条件下可能存在一定误差。在今后的研究中,将进一步讨论在复杂背景条件下的农作物病害图像分割问题。

参考文献:

[1] Navrozidis I, Alexandridis T K, Dimitrakos A, et al. Identification of Purple Spot Disease on Asparagus Crops Across Spatial and Spectral Scales[J]. Computers and Electronics in Agriculture, 2018, 148: 322-329.

[2] Ma J, Du K, Zheng F, et al. A Recognition Method for Cucumber Diseases Using Leaf Symptom Images Based on Deep Convolutional Neural Network[J]. Computers and Electronics in Agriculture, 2018, 154: 18-24.

[3] 胡新,许艳丽,李春杰. 利用作物多样性控制病害研究进展[J]. 农业系统科学与综合研究, 2011, 27(1): 118-122.

[4] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks[J]. Communications of the ACM, 2017, 60(6): 84-90.

[5] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[C]//3rd International Conference on Learning Representations, San Diego, CA, USA, 2015: arxiv.1409.1556

[6] Szagedy C, Vanhoucke V, Ioffe S, et al. Rethinking the Inception Architecture for Computer Vision[C]//2016 IEEE International Conference on Computer Vision (IC-

CV),Las Vegas,NV,USA,2016:2818-2826.
[7] He K,Zhang X,Ren S,et al. Deep Residual Learning for Image Recognition[C]//2016 IEEE International Conference on Computer Vision(ICCV),Las Vegas,NV,USA,2016:770-778.
[8] 李好,邱卫根,张立臣. 改进 ShuffleNet V2 的轻量级农作物病害识别方法[J]. 计算机工程与应用,2022,58(12):260-268.
[9] 王焕鑫,沈志豪,刘泉,等. 基于改进 MobileNetV2 模型的农作物叶片病害识别研究[J]. 河南农业科学,2023,52(4):143-151.
[10] Badrinarayanan V,Kendall A,Cipolla R. Segnet:A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(12):2481-2495.
[11] Ronneberger O,Fischer P,Brox T. U-Net:Convolutional Networks for Biomedical Image Segmentation[C]//18th International Conference on Medical Image Computing and Computer-Assisted Intervention,Munich,Germany,2015:234-241.
[12] Zhao H S,Shi J P,Qi X J,et al. Pyramid Scene Parsing Network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR),Honolulu,HI,USA,2017:6230-6239.
[13] Chen L C,Zhu Y K,Papandreou G,et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation[C]//15th European Conference on Computer Vision(ECCV),Munich,Germany,2018:833-851.
[14] 王守印,陈健,万佳泽,等. 基于跨层非局部融合和 DeepLabV3+的 PCB 图像分割算法[J]. 电子器件,2024,47(6):1547-1555.
[15] 钟昌源,胡泽林,李淼,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.
[16] 赵小虎,李晓,叶圣,等. 基于改进 U-Net 网络的多尺度番茄病害分割算法[J]. 计算机工程与应用,2022,58(10):216-223.
[17] Ferentinos K P. Deep Learning Models for Plant Disease Detection and Diagnosis[J]. Computers and Electronics in Agriculture,2018,145:311-318.
[18] Yu F,Koltun V. Multi-Scale Context Aggregation by Dilated Convolutions[C]//4th International Conference on Learning Representation,San Juan,Puerto Rico,2016:arxiv.1511.07122
[19] Chen L C,Papandreou G,Kokkinos I,et al. Deeplab:Semantic Image Segmentation with Deep Convolutional Nets,Atrous Convolution,and Fully Connected Crfs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,40(4):834-848.
[20] Cao Y,Xu J R,Lin S,et al. GCNet:Non-Local Networks Meet Squeeze-Excitation Networks and Beyond[C]//2019 IEEE/CVF International Conference on Computer Vision Workshops(ICCV),Seoul,South Korea,2019:1971-1980.
[21] Ho Y,Wookey S. The Real-World-Weight Cross-Entropy Loss Function:Modeling the Costs of Mislabeling[J]. IEEE Access,2019,8:4806-4813.
[22] Li X Y,Sun X F,Meng Y X,et al. Dice Loss for Data-Imbalanced NLP Tasks 58th Annual Meeting of the Association for Computationa Linguistics,Online,2020:465-476.
[23] Kingma D P,Ba J. Adam:A Method for Stochastic Optimization[EB/OL]. (2017-01-30).
[24] Long J,Shelhamer E,Darrell T. Fully Convolutional Networks for Semantic Segmentation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition,Boston,MA,USA,2015:3431-3440.
[25] Yuan Y,Chen X,Wang J. Object-Contextual Representations for Semantic Segmentation[C]//16th European Conference on Computer Vision,Glasgow,UK,2020:173-190.
[26] Strudel R,Garcia R,Laptev I,et al. Segmenter:Transformer for Semantic Segmentation[C]//2021 IEEE/CVF International Conference on Computer Vision,Montreal,QC,Canada,2021:7242-7252.
[27] Zhang W W,Pang J M,Chen K,et al. K-Net:Towards United Image Segmentation[C]//35th International Conference on Neural Information Processing Systems,Online,2021:10326-10338.
[28] Oquab M,Bottou L,Laptev I,et al. Learning and Transferring Mid-Level Image Representations Using Convolutional Neural Networks[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition(CVPR),Columbus,OH,USA,2014:1717-1724.

刘伟立(1998—),男,硕士,主要从事深度学习与图像识别研究;

惠国华(1980—),男,教授,主要从事智慧农业领域的研究,deliver1982@163.com。