doi:10.6041/j.issn.1000-1298.2026.22.032

层次特征与语义融合解码的苹果叶片病害语义分割

邬开俊[1]　汪　滢[1]　张治瑞[1]　陈融禹[1]　张红娜[2]
(1.兰州交通大学电子与信息工程学院,兰州 730070;2.内蒙古民族大学物理与电子信息学院,通辽 028000)

摘要:苹果叶片病害语义分割是智能农业病害监测的重要任务。针对自然场景中病害分割病斑尺度变化大、边界模糊、复杂背景干扰的挑战,本文提出层次特征与语义融合解码的苹果叶片病害语义分割模型。首先,提出层次特征网络(Hierarchical feature network,HFN),通过五阶段特征提取和残差结构实现从边缘轮廓到全局语义的建模。其次,为增强对病斑尺度变化和边缘结构的适应能力,设计边缘采样空洞空间金字塔模块(Edge sampling atrous spatial pyramid,ES-ASPP),捕捉多尺度语义信息并细化边缘结构。最后,由于语义分割任务中需要高效解码器以还原空间信息,因此提出语义融合解码器(Semantic fusion decoder,SFD),协同表达浅层细节与深层语义,提升边界还原与小病斑检测能力。试验结果表明,本文所提方法在苹果叶片病害数据集上的 mIoU 较基线模型提升 5.7 个百分点,参数量下降 6.2%;且在 PlantSeg 数据集泛化试验中相较 DeepLabV3Plus mIoU 提升 4.3 个百分点,具备良好的分割性能,适用于自然场景中农业病害分割任务。

关键词:语义分割;苹果叶片;病害检测;编解码器;骨干网络;多尺度特征
中图分类号:TP391.4;S431.9　　**文献标识码**:A　　**文章编号**:1000-1298(2026)22-0369-09　　**OSID**:

Hierarchical Feature and Semantic Fusion Decoding for Semantic Segmentation of Apple Leaf Disease

Wu Kaijun[1]　Wang Ying[1]　Zhang Zhirui[1]　Chen Rongyu[1]　Zhang Hongna[2]
(1. *School of Electronic and Information Engineering*, *Lanzhou Jiaotong University*, *Lanzhou* 730070, *China*
2. *School of Physics and Electronic Information*, *Inner Mongolia University for Nationalities*, *Tongliao* 028000, *China*)

Abstract: Semantic segmentation of apple leaf disease is an important task for disease monitoring in smart agriculture. Aiming at the challenges of large scale changes, blurred boundaries, and complex background interference of disease segmentation spots in natural scenes, a semantic segmentation model for apple leaf disease with hierarchical feature and semantic fusion decoding was proposed. Firstly, hierarchical feature network (HFN) was proposed to achieve modelling from edge contour to global semantics through five-stage feature extraction and residual structure. Secondly, in order to enhance the adaptability to spot scale variations and edge structure, the edge sampling atrous spatial pyramid (ES-ASPP) module was designed to capture multi-scale semantic information and refine the edge structure. Finally, since an efficient decoder was required in the semantic segmentation task to restore spatial information, semantic fusion decoder (SFD) was proposed to synergistically express the shallow details and deep semantics to enhance the boundary restoration and small lesion detection. The experimental results showed that the mIoU of the proposed method on the apple leaf disease dataset was improved by 5.7 percentage points compared with that of the baseline model, and the number of parameters was decreased by 6.2%; and the mIoU was improved by 4.3 percentage points in the generalisation experiment on the PlantSeg dataset, which can provide a good segmentation performance, and it was suitable for the task of segmentation of agricultural diseases in natural scenes.

Key words: semantic segmentation; apple leaves; disease detection; encoder-decoder; backbone network; multi-scale features

收稿日期:2025-06-27　修回日期:2025-07-18
基金项目:甘肃省自然科学基金项目(23JRRA913)和内蒙古自治区高校基本科研业务费项目(GXKY25Z012)
作者简介:邬开俊(1978—),男,教授,博士,主要从事深度学习和计算机视觉研究,E-mail:wkj@mail.lzjtu.cn

0 引言

苹果因高营养价值深受青睐,但其叶片病害等导致苹果产量损失,快速检测并精准施药对保障苹果健康生长至关重要[1]。然而,田间环境下病斑微小、边界模糊及病斑尺度变化大等挑战使传统人工目视和早期分割方法效果有限[2]。

近年来,深度学习在叶片病害语义分割任务中取得显著进展。Storey 等[3]采用 ResNet-50[4]和 MobileNetV3[5]作为 Mask-RCNN[6]的骨干网络,验证其在苹果叶片病害图像分割中的适用性。Agarwal 等[7]以 ResNet 为 U-Net[8]的编码器,实现对农作物叶片病害的高效分割,展现深度卷积网络在病害特征提取方面的潜力。Li 等[9]基于 DeepLabV3Plus[10]使用 ResNet-50 骨干网络解决叶片病害分割中目标尺度变化和环境因素干扰等问题。以上方法主要依赖于传统的卷积结构并基于骨干网络进行改进,对病害的局部纹理和形状特征提取有效,但在建模远距离依赖关系和全局语义方面存在局限,难以应对微小病斑检测和尺度多样化的问题。为提升对病斑尺度变化的适应性,多尺度方法逐渐应用于病害分割任务,以兼顾大范围病斑的上下文建模和微小病斑定位。鲍文霞等[11]设计多尺度卷积路径以提升对不同尺度病斑的感知能力。Zeng 等[12]通过分组卷积与多分支空洞卷积的协同设计,以捕捉橡胶树叶病斑的多尺度特征。然而,这些方法对边界模糊和复杂边缘的刻画能力有限。此外,高效解码器对病斑边界与形态的准确重建至关重要。邹一波等[13]结合特征距离和空间距离的双近邻插值提升边界还原精度。Ding 等[14]在解码器中采用密集连接以提升特征重用与细节恢复能力。但在复杂背景和微小病斑检测方面依然存在误分与漏分现象。

针对上述不足,本文提出层次特征与语义融合解码的苹果叶片病害语义分割模型,在 DeepLabV3Plus 基础上进行 3 项关键改进:①构建基于 ResNet-50 的 HFN 骨干网络,通过融合多尺度卷积和 Transformer[15]结构,以增强微小病斑的识别能力并提升全局语义建模效果。②引入 ES-ASPP 并与 HFN 构成编码器,ES-ASPP 采用多尺度空洞卷积和边缘采样注意力机制,有效应对病斑大小不一与边界模糊问题,提升结构刻画能力。③设计 SFD,逐级融合不同层次特征并引入注意力门控,以期实现病斑语义还原与边缘细节重建,提升微小病斑检测与复杂背景下分割精度。

1 网络模型

为应对苹果叶片病害语义分割中所面临的多重挑战,提出层次特征与语义融合解码的苹果叶片病害语义分割模型,如图 1 所示。网络由 HFN、ES-ASPP 和 SFD 3 个部分组成。输入图像经 HFN 提取低、中、高层次的多尺度特征,高层次特征进一步输入 ES-ASPP 进行上下文增强;增强后特征与 HFN 第 1 阶段输出的低层次特征及第 2 阶段输出的中层次特征在 SFD 中逐级上采样时融合,最终生成分割结果。该分阶段特征交互策略有效整合局部细节与全局语义,实现多尺度特征的有效互补与优化。HFN、ES-ASPP 与 SFD 的协同工作,在苹果叶片病害语义分割任务中展现出较高的分割精度与良好的适应性。

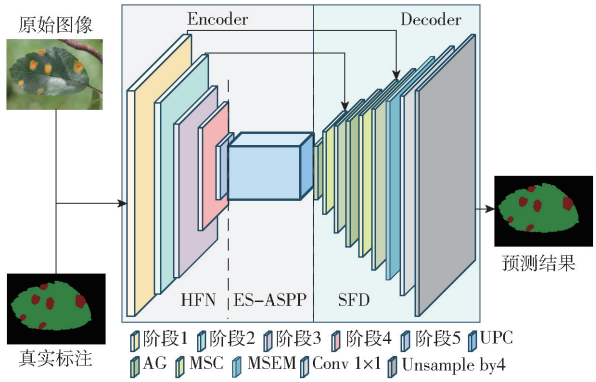


图 1 网络模型整体框架

Fig. 1 Overall framework of network

1.1 层次特征网络

针对苹果叶片病斑尺度差异和微小病斑难以检测的问题,提出融合 CNN 与 Transformer 结构的 HFN 骨干网络,以实现多尺度、高语义一致性的病斑区域建模。如图 2 所示,网络由 5 个阶段组成,从上到下依次提取边缘轮廓、病斑边界与显著响应特征、多尺度纹理与结构信息、长距离空间依赖关系与跨区域语义一致性,以及复杂背景下的全局上下文语义,逐步构建从局部细节到整体结构的特征表达。HFN 结合局部建模的精细性与全局感知的鲁棒性,提升模型对复杂病斑形态的适应能力与分割精度。

阶段 1 采用 7×7 卷积扩大感受野,捕捉初始轮廓边缘、色域维度等底层特征,匹配病斑的色彩及纹理差异。阶段 2 通过连续 3×3 卷积压缩特征,并结合 SE(Squeeze-and-excitation)通道注意力[16]增强对病斑边界、叶脉断裂等高频区域的感知能力,且以深度可分离卷积控制参数规模。针对病斑区域边缘模糊、纹理复杂问题,引入残差式增强结构,采用逐元素先乘后加的融合方式替代传统残差相加,在保留原始语义信息的同时显式强化边缘响应,提升微小

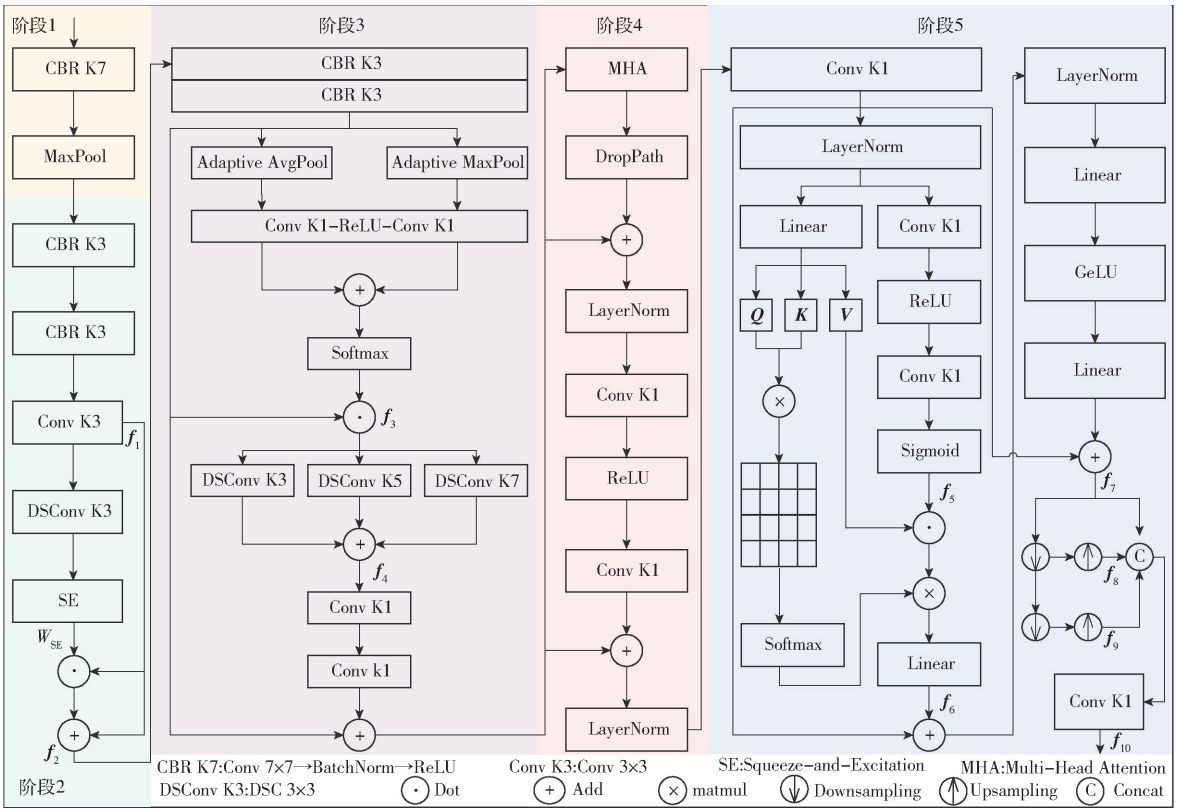


图 2　层次特征网络

Fig. 2　Hierarchical feature network

病斑及细节刻画能力，即

$$\boldsymbol{f}_2 = \boldsymbol{f}_1 + \boldsymbol{f}_1 \odot \boldsymbol{W}_{SE} \tag{1}$$

式中　$\boldsymbol{f}_1$——阶段 2 中由卷积层提取的中间特征

$\boldsymbol{W}_{SE}$——SE 模块生成的通道注意力权重

$\boldsymbol{f}_2$——经过 SE 注意力重标定和残差融合后的输出特征

$\odot$——逐元素乘法

阶段 3 通过 3×3 卷积实现空间压缩与通道拓展，强化病斑区域的上下文建模与全局感知能力。通过并行自适应平均池化与最大池化提取每个通道的全局统计，生成 2 种互补的通道描述，随后经共享的 2 层 1×1 卷积与 ReLU 进行非线性变换与重标定。在通道维度采用 Softmax 引导的权重学习机制实现高响应区域的特征聚焦，抑制背景干扰，提升病斑判别性。空间维度通过卷积核尺寸为 3×3、5×5 和 7×7 的多尺度深度可分离卷积构建并行路径，以捕捉不同感受野下的纹理与结构信息，适应病斑尺度与形态变化，即

$$\boldsymbol{f}_4 = \sum_{i=3,5,7} \mathrm{DSC}_{i\times i}(\boldsymbol{f}_3) \tag{2}$$

式中　$\boldsymbol{f}_3$——经通道注意力加权后得到的特征图

$\mathrm{DSC}_{i\times i}(\cdot)$——采用 $i\times i$ 卷积核的深度可分离卷积运算

i——取值集合为{3,5,7}，对应 3 个不同感受野的并行卷积分支

$\boldsymbol{f}_4$——3 个尺度分支输出经逐元素求和后得到的多尺度融合特征图

该结构结合残差连接，保留原始语义信息并促进梯度传播，为后续分支提供高语义特征支持。

阶段 4 首先将二维特征图展平为序列后输入多头自注意力机制（Multi-head attention，MHA），捕捉远距离像素的上下文关系，强化病斑间的全局语义一致性与结构建模，提升对离散分布病斑的适应性。随后，通过逐点卷积实现通道非线性组合，增强对病斑纹理与结构感知。在训练阶段引入 DropPath[17] 随机抑制部分残差路径，降低特征通道依赖并提升训练稳定性。此阶段通过双分支残差结构维持语义连续与跨层一致性，实现深层语义融合，为边缘模糊、尺度不一的病斑提供更强的全局判别支持。

阶段 5 首先通过 1×1 卷积压缩通道，将高维特征投影至低维语义空间以降低计算量并保留关键信息；随后展平为序列并经层归一化输入自注意力机制，通过线性变换生成查询 $\boldsymbol{Q}$、键 $\boldsymbol{K}$ 和值 $\boldsymbol{V}$ 向量，查询与键向量之间进行矩阵乘法操作，生成空间注意力图以建模远程依赖，同时利用 2 层逐点卷积与 Sigmoid 生成权重图 $\boldsymbol{f}_5$，用于调节每个空间位置的表示强度。对注意力值 $\boldsymbol{V}$ 进行显式加权，突出病斑响应并抑制背景干扰，即

$$head_i = S\left(\frac{\boldsymbol{Q}_i \boldsymbol{K}_i^{\mathrm{T}}}{\sqrt{d_k}}\right)(\boldsymbol{V}_i \odot \boldsymbol{f}_5) \quad (i = 1,2,\cdots,h) \tag{3}$$

$$\boldsymbol{f}_6 = L(\mathrm{Concat}(head_1, head_2, \cdots, head_h)) \tag{4}$$

式中 $\boldsymbol{Q}_i$、$\boldsymbol{K}_i$、$\boldsymbol{V}_i$——第 i 个注意力头查询、键和值特征

d_k——键向量维度

h——注意力头数量

S——Softmax 函数

$head_i$——第 i 个注意力头输出

Concat(·)——通道拼接

L——线性映射

$\boldsymbol{f}_6$——多头注意力融合后输出特征

注意力输出后经前馈子网络采用 2 层逐点卷积与非线性激活实现通道维度的非线性映射与特征交互，进一步强化病斑结构的内部响应差异。整体结构中引入残差连接与层归一化，以保证训练稳定性与深层信息的一致性传递。为增强模型对病斑尺度适应性，模块末端采用 2 级 3×3 卷积下采样获取不同感受野特征，再通过双线性插值将其统一还原至原始分辨率。最终，3 路尺度特征在通道维度进行拼接，并通过 1×1 卷积融合形成统一高维特征输出，实现病斑区域局部与全局、多尺度的综合语义整合，即

$$\boldsymbol{f}_{10} = \mathrm{Conv}_{1\times1}(\mathrm{Concat}[\boldsymbol{f}_7, \boldsymbol{f}_8, \boldsymbol{f}_9]) \tag{5}$$

式中 $\boldsymbol{f}_7$——原尺度分支输出特征

$\boldsymbol{f}_8$、$\boldsymbol{f}_9$——不同下采样尺度分支经上采样恢复后特征

$\mathrm{Conv}_{1\times1}(\cdot)$——1×1 卷积融合操作

$\boldsymbol{f}_{10}$——阶段 5 输出的多尺度融合特征

1.2 边缘采样空洞空间金字塔模块

ES-ASPP 集成多尺度卷积模块（Multi-scale convolution module，MCM）和边缘采样注意力模块（Edge sampling attention module，ESAM），以提升模型对病斑区域的尺度变化适应性与边界结构刻画能力，如图 3 所示。

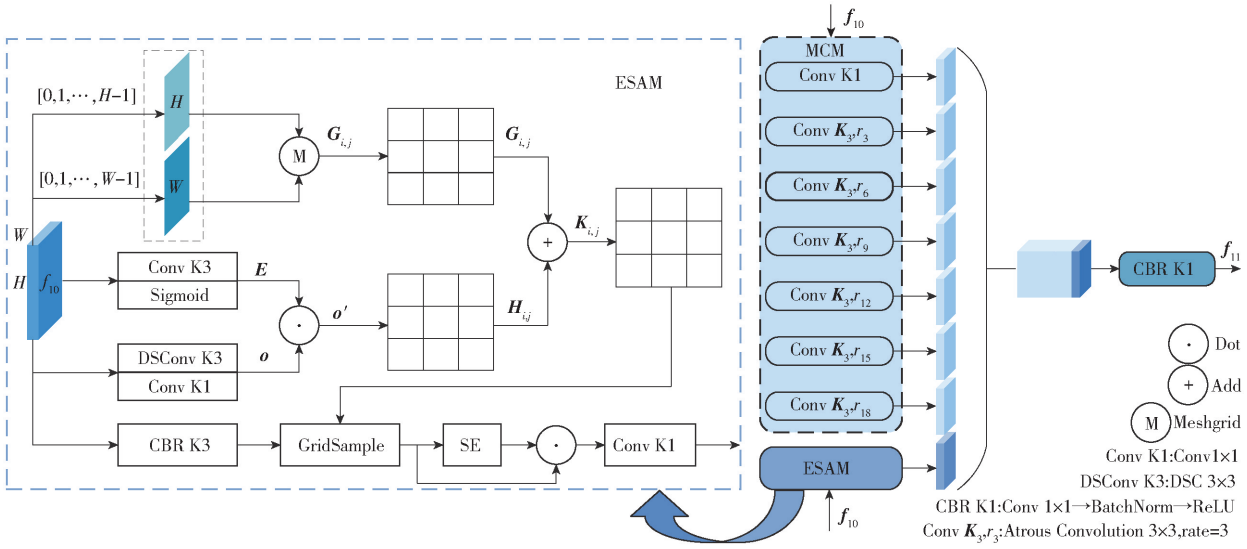


图 3 边缘采样空洞空间金字塔模块

Fig. 3 Edge sampling atrous spatial pyramid module

HFN 的输出特征 $\boldsymbol{f}_{10}$ 同时进入 MCM 和 ESAM。MCM 构建 7 组并行的空洞卷积路径（膨胀率为 1、3、6、9、12、15、18），以多尺度感受场模拟局部至全局语义，增强对病斑尺度的适应性。与此同时，ESAM 引入边缘引导的自适应采样机制缓解尺度变化和复杂背景引起的边界模糊。ESAM 首先使用 3×3 卷积与 Sigmoid 生成边缘响应图 $\boldsymbol{E}$，该边缘图表征图像中潜在的结构轮廓信息，并作为权重参与调制偏移场的预测过程。偏移分支由深度可分离卷积和点卷积组成，输出偏移张量 $\boldsymbol{o}$，表示每个空间位置上采样窗口内的二维偏移量。偏移通过边缘图进行调制，得到边缘感知的偏移张量 $\boldsymbol{o}'$，即

$$\boldsymbol{o}' = \boldsymbol{o} \odot \boldsymbol{E} \tag{6}$$

使偏移预测对边缘区域更敏感，增强结构感知能力。对每个位置对应的 K^2 个采样点偏移向量求和，整合成 1 个总偏移向量 $\boldsymbol{H}_{i,j}$，即

$$\boldsymbol{H}_{i,j} = \sum_{p=1}^{K^2} \boldsymbol{o}'^{(p)}_{i,j} \tag{7}$$

式中 $\boldsymbol{o}'^{(p)}_{i,j}$——位置 (i,j) 处第 p 个采样点对应的边缘感知偏移向量

基于输入特征图 $\boldsymbol{f}_{10}$ 高 H 和宽 W，构造归一化坐标网格 $\boldsymbol{G}_{i,j}$，即

$$\boldsymbol{G}_{i,j} = \left(\frac{2j}{W} - 1, \frac{2i}{H} - 1\right) \tag{8}$$

其中 $i \in [0, H-1]$ $j \in [0, W-1]$

式中，i、j 分别表示特征图的行、列索引。

然后将调制后的偏移张量映射到坐标空间，与

基础网格相加形成结构适应的采样网格 $\boldsymbol{K}_{i,j}$，即

$$\boldsymbol{K}_{i,j}=\boldsymbol{G}_{i,j}+\boldsymbol{H}_{i,j} \tag{9}$$

使用 3×3 卷积提取主干特征，并通过网格采样操作基于结构感知网格，对主干特征进行空间变形采样，从而增强对不规则结构和边缘细节的建模能力。变形采样后的特征进一步输入通道注意力 SE 分支，生成通道权重系数以调节特征响应。最终，通过逐通道加权与 1×1 卷积映射输出结果，实现空间几何建模与通道语义选择的深度融合。所有输出路径在通道维度上进行拼接，并通过 1×1 卷积压缩通道数以形成统一表征，结合归一化、非线性激活与 Dropout 操作，提升特征表达能力与泛化性。

1.3　语义融合解码器

在苹果叶片病害语义分割中，SFD 解码器负责多层特征的语义还原与空间细化。如图 4 所示，其以编码器中 ES－ASPP 输出的高层语义特征 $\boldsymbol{f}_{11}$ 为主干，逐级融合 HFN 第 2 阶段的中层特征与第 1 阶段的低层特征，实现病斑区域形态与边界的渐进式还原。具体而言，解码器首先通过 1×1 卷积压缩并对齐中、低层特征通道；主干特征经上采样卷积（Upsampling conv，UPC）模块增强上下文信息，并在注意力门控（Attention gate，AG）引导下融合中层结构特征，随后与中层特征通道拼接后输入多尺度卷积（Multi-scale conv，MSC）模块进行多尺度建模，适应病斑结构异质性与尺度变化。继而采用相同策略引入低层边缘细节，进一步细化病斑边缘、叶脉断裂等高频区域，提升空间定位精度。融合特征经 MSEM 模块（与 HFN 第 3 阶段去除 2 层 3×3 卷积－BN－ReLU 的结构相同）增强语义响应并抑制背景，最终通过 1×1 卷积生成类别预测图。该设计结合多尺度建模与通道注意力协同作用，实现从语义到结构的逐步重建，有效提升微小病斑检测、不规则分布识别及复杂背景下的辨识能力。

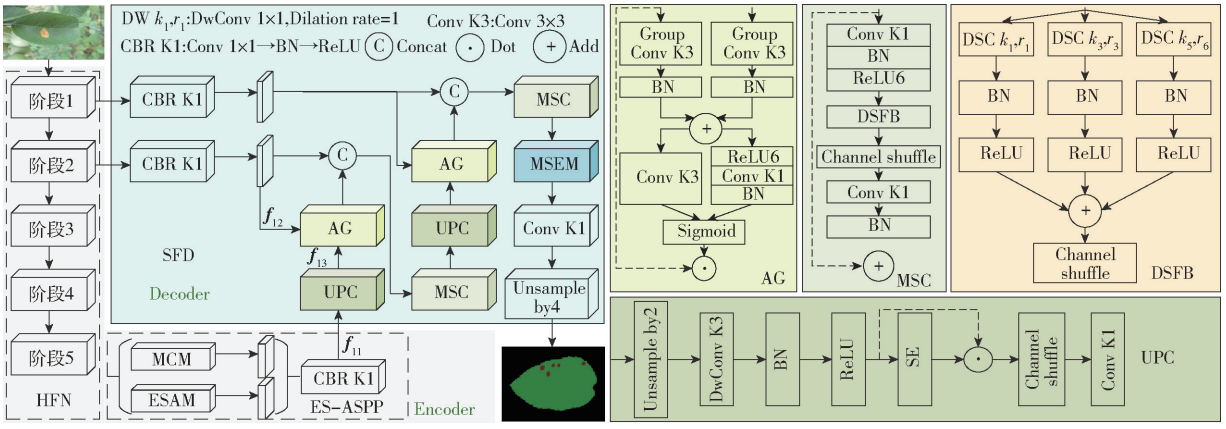


图 4　语义融合解码器结构

Fig. 4　Semantic fusion decoder structure

UPC 模块作为解码初始阶段的关键结构，旨在增强上采样特征的上下文语义与通道表达能力。其首先结合双线性插值与深度可分离卷积构建轻量化上采样路径，在提升空间分辨率的同时保留关键语义信息；随后引入 SE 通道注意力自适应调节通道响应，强化病斑特征并抑制背景干扰。为进一步增强通道间交互与特征表达多样性，UPC 引入通道混洗（Channel shuffle）[18] 策略，有效打破卷积间的信息隔离，促进跨通道信息融合。最终，通过逐点卷积实现通道压缩与映射，将上采样后的高维特征转换为统一表示，以便后续解码分支处理；当输入输出通道一致时，模块引入残差连接以保持语义一致性并稳定梯度传播。

AG 模块用于融合全局与局部特征，增强病斑显著性表达并刻画边缘细节。首先，将局部特征 $\boldsymbol{f}_{12}$ 和全局特征 $\boldsymbol{f}_{13}$ 通过组卷积与批归一化把通道数映射到一致维度，缓解特征分布偏差。并相加生成融合特征，兼具上下文语义与细粒度信息。随后，融合特征进入两分支，边缘引导分支经 3×3 卷积和 Sigmoid 生成边缘权重图，突出病斑边界；通道引导分支通过非线性激活、点卷积和批归一化提取通道注意力，强化判别性通道响应。最终，将局部特征 $\boldsymbol{f}_{12}$ 同时与通道引导权重和边缘权重逐像素相乘，实现空间和通道的双重加权，有效增强病斑特征表达并抑制背景干扰，提升分割精度和边缘清晰度。

MSC 模块旨在实现多尺度感知、结构建模与语义重组。输入特征经点卷积升维，扩展通道以丰富表达；随后特征被送入内部嵌套的深度卷积融合混洗块（Depthwise fusion shuffle block，DFSB），该结构并行构建多分支深度可分离卷积，每个分支对应一组特定的卷积核尺寸与空洞率（1×1 卷积对应 $d=1$，3×3 卷积对应 $d=3$，5×5 卷积对应 $d=6$），在保

持计算效率的同时建模特定尺度的上下文结构,小感受野卷积强化病斑局部边缘与纹理响应,而大感受野卷积捕捉病斑的全局分布与远程依赖,形成尺度多样、语义互补的特征集合。多尺度特征通过通道拼接整合所有分支输出,并引入通道混洗机制打破组间隔离,增强跨通道特征交互。融合后的特征再经 1×1 卷积映射至目标维度,形成统一的输出表征。输入输出通道一致时加入残差连接,以缓解深层信息退化与梯度消散问题,同时增强边缘细节的保留能力。整体而言,MSC 模块通过构建多尺度卷积、特征融合与残差增强,有效提升模型对病斑纹理、边界过渡和尺度变化的感知能力,是语义解码过程中的关键结构。

2 试验

2.1 数据集

为真实反映苹果叶片病害自然形态,自行采集并构建一套苹果叶片病害语义分割数据集。数据集采集工作在甘肃省天水市苹果园完成,采集过程中未对叶片进行任何人为干预,保持自然生长状态。图像采集后使用 labelme 工具进行像素级精细标注,确保标签的准确性与完整性。数据集中包含的病害叶片涵盖多种典型病害类型,包括苹果锈病、褐斑病、黑点病。在数据标注过程,并非单纯以病害类型作为唯一分割目标,而是重点关注病害区域病斑形态、边缘轮廓及其在叶片上分布特征。自建数据集样例如图 5 所示,图 5a 展示了不同尺度病斑的典型样本,其中蓝色框表示小尺度病斑区域,红色框表示大尺度病斑区域。受病斑尺度变化大、病斑边缘模糊、病斑过小及复杂背景等因素的影响,标注工作与病斑分割难度大。数据集划分为病害类和叶片类。这种标注策略旨在为模型提供对病害区域更全面的感知能力,支持更广泛的任务场景,如病害严重程度评估、病斑边界变化检测等。最终构建的数据集共包含 2 000 幅图像,按比例 7∶2∶1划分为训练集(1 400 幅)、验证集(400 幅)和测试集(200 幅)。



(a) 病斑尺度变化大 (b) 病斑边缘模糊 (c) 微小病斑 (d) 复杂背景

图 5 自建数据集样例

Fig. 5 Sample self-constructed dataset

为进一步验证本文所提模型在不同植物病害场景下通用性与鲁棒性,还使用公开植物病害数据集 PlantSeg[19],包含 115 种不同植物疾病的图像。为实现与本文自建苹果叶片病害数据集在任务层面的功能对齐,并未使用 PlantSeg 原始的多类标签结构,而是将所有病害类别统一合并为“病害类”,其余区域视为“背景类”,从而构建了 1 个与自建数据集一致的二分类语义分割任务。该设置使模型能够专注于病害区域的提取能力评估,而非具体病种识别,有助于统一训练目标,提升在不同数据源间的泛化表现。同时,该策略也更契合病斑检测、病害预警等前端部署任务的应用需求。

2.2 试验设置及评价指标

试验环境为:Windows 11 系统,Python 3.8,使用显卡为 NVIDIA GeForce RTX 4060,整个环境基于 PaddlePaddle 框架。优化器选择 SGD(动量为 0.9)优化器;学习率为多项式衰减,初值为 0.01;权重衰减率为 4×10^{-5};共迭代 5 000 次。图像增强策略选用固定尺寸调整到 512 像素×512 像素、随机水平翻转(概率 0.5)、随机垂直翻转(概率 0.1)、以固定步长随机缩放图像(最小缩放因子 0.75、最大缩放因子 1.25、缩放步长 0.25)、随机剪裁、随机高斯噪声(添加概率 0.5、噪声标准差最大值 10.0)、模糊(概率 0.1、高斯模糊、中值模糊、均值模糊)、随机旋转、随机仿射变换(平移、旋转、缩放)、归一化(均值 0.5、标准差 0.5),随机调整亮度、对比度、饱和度、色调和锐度。

采用平均交并比(mIoU)、整体像素精度(Accuracy)和模型参数量(Params)评估模型性能。其中,mIoU 用于衡量模型预测分割结果与真实标签之间的重叠程度,mIoU 越高,表示模型分割精度越高。Accuracy 表示正确分类像素占总像素比例。

2.3 不同模型性能测试

本文模型与 6 种经典语义分割模型进行测试,如表 1 所示,在 mIoU(91.3%)上较基线模型 DeepLabV3Plus(ResNet-50)提升 5.7 个百分点,参数量减少 6.2%,表明所提方法在苹果叶片病害分割任务中的优越性。可视化结果如图 6 所示,本文模型在多个场景中均表现出更强的分割能力,在图

像 1 中，能够完整分割大病斑并准确检测 6 个微小病斑，无漏检现象；在图像 2 中，成功分割出 6 个病斑且未发生类别误判；在图像 3 中，对叶片与病斑区域的边界识别更为完整准确。这一性能提升得益于所提模型中各模块协同作用，HFN 通过局部细节与全局语义的分阶段交互，增强了模型对病斑形态差异与微小病斑的适应能力；ES－ASPP 引入多尺度卷积与边缘感知采样机制，提升了对病斑尺度变化与边缘模糊区域的建模能力，有效减少误分割与漏分割现象；SFD 解码器结合多层次特征与注意力机制，强化病斑区域的语义一致性和边缘定位精度，实现更精细的空间还原与结构重建。

表 1　不同模型在自建苹果叶片病害数据集上性能测试结果

Tab. 1　Performance result of different models on self-built apple leaf disease dataset

模型	平均交并比/%	准确率/%	参数量
DeepLabV3Plus(ResNet－50)	85.43	97.00	2.679×10^7
PSPNet (ResNet101)[20]	80.19	95.43	6.700×10^7
SegNext－T[21]	87.53	97.65	4.300×10^6
SeaFormer－B[22]	86.02	97.19	8.600×10^6
AISOA－SSformer[23]	81.59	95.85	1.471×10^7
SCTNet－B[24]	88.65	97.98	1.740×10^7
本文模型	91.13	99.60	2.512×10^7

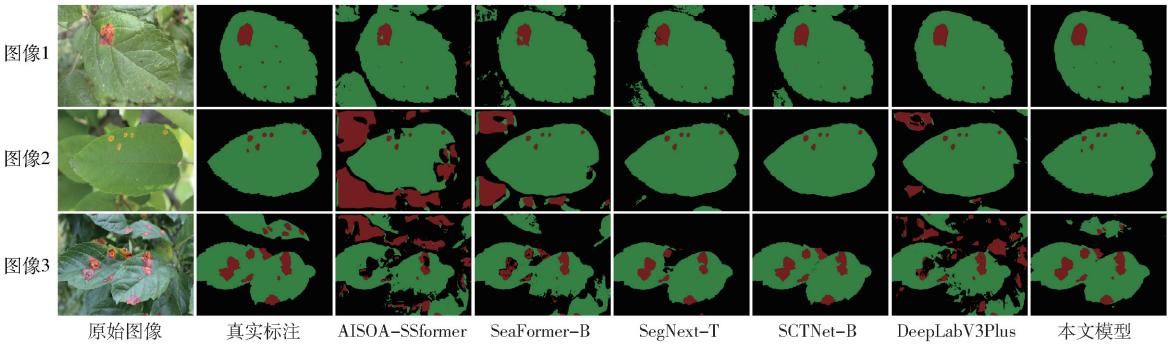


图 6　自建苹果叶片病害数据集可视化效果

Fig. 6　Visualization effects of self-built apple leaf disease dataset

2.4　消融试验

2.4.1　模块消融试验

模块消融试验旨在分别验证所引入的 HFN 骨干网络、ES－ASPP 模块及 SFD 解码器对语义分割性能的提升效果。为评估 HFN 的特征提取能力，将 DeepLabV3Plus 中 ResNet－50 的初始阶段及其后 4 个阶段依次替换为 HFN 对应层次。结果如表 2 所示，阶段 1 至阶段 5 的逐步替换，mIoU 提升 2.8 个百分点，得益于 CNN 与 Transformer 的融合设计，HFN 能够分层提取边缘、纹理、结构与语义信息，实现局部与全局特征的有效协同，提升模型对病斑形态变化及微小病斑的适应能力。

表 2　HFN 骨干网络分析

Tab. 2　HFN backbone network analysis

配置	阶段 1	阶段 2	阶段 3	阶段 4	阶段 5	平均交并比/%	准确率/%
基线模型(DeepLabV3Plus/ResNet－50)	×	×	×	×	×	85.43	97.00
替换阶段 1	√	×	×	×	×	85.63	97.10
替换阶段 1－2	√	√	×	×	×	85.93	97.15
替换阶段 1－3	√	√	√	×	×	86.93	97.42
替换阶段 1－4	√	√	√	√	×	87.33	97.63
完整 HFN	√	√	√	√	√	88.23	97.85

注：√表示采用此阶段，×表示不采用此阶段。下同。

将原始 ASPP 中的空洞卷积替换为 MCM 模块，并将最大池化替换为 ESAM 模块。结果如表 3 所示，mIoU 分别提高 0.69、1.6 个百分点，验证了 MCM 通过 7 组空洞卷积能够增强多尺度语义表达，强化模型对病斑大小变化的适应能力；ESAM 提升了对边缘敏感区域的关注能力，有效缓解边缘模糊引发的误检和漏检。

表 3　ES－ASPP 模块分析

Tab. 3　ES－ASPP module analysis

配置	MCM	ESAM	平均交并比/%	准确率/%
基线模型(DeepLabV3Plus)	×	×	85.43	97.00
基线模型＋MCM	√	×	86.12	97.22
基线模型＋MCM＋ESAM	√	√	87.72	97.71

对 SFD 解码器各子模块进行逐项消融，结果如表 4 所示。UPC 替换原始上采样操作后，mIoU 提升 0.65 个百分点，高效恢复空间分辨率并保留特征表达。AG 明显提升边缘与前景分离能力，mIoU 提高 0.89 个百分点。MSC 使用并行多空洞率分支与通道混洗，建模病斑纹理、边界、形态变化，mIoU 提高 0.08 个百分点。MSEM 抑制背景噪声干扰，mIoU 提高 0.6 个百分点。完整 SFD 的 mIoU 相比基线模型提高 2.22 个百分点，表明该模块在重建语义及细化边界方面表现突出。

表 4 SFD 解码器分析

Tab. 4 SFD decoder analysis

配置	UPC	AG	MSC	MSEM	平均交并比/%	准确率/%
基线模型(DeepLabV3Plus)	×	×	×	×	85.43	97.00
基线模型 + UPC	√	×	×	×	86.08	97.21
基线模型 + UPC + AG	√	√	×	×	86.97	97.48
基线模型 + UPC + AG + MSC	√	√	√	×	87.05	97.50
基线模型 + UPC + AG + MSC + MSEM	√	√	√	√	87.65	97.68

2.4.2 整体消融试验

为验证各模块组合对模型整体性能的影响，设计逐步添加的组合试验，评估模型性能在引入 HFN、ES－ASPP、SFD 后的逐级变化情况。结果如表 5 所示，随着模块逐步引入，mIoU 提至 91.13%。结果表明，所提模块在整体架构中能够协同工作，通过局部细节与全局语义交互、多尺度与边缘信息增强、特征还原与空间重建等机制综合提升病斑分割性能，有效增强模型对病斑尺度变化、边界模糊及复杂背景干扰的适应能力。

表 5 不同模块消融试验结果

Tab. 5 Ablation results of different modules

HFN	ES－ASPP	SFD	平均交并比/%	准确率/%
√	×	×	88.23	97.85
×	√	×	87.72	97.71
×	×	√	87.65	97.68
√	√	×	89.63	98.53
√	×	√	89.23	98.20
×	√	√	88.96	98.39
√	√	√	91.13	99.60

2.5 泛化试验

尽管所提模型在自建苹果叶片病害数据集上表现较好，但其泛化能力仍需进一步验证。为此，引入公开植物病害分割数据集 PlantSeg 进行评估，以验证模型在不同植物种类下的适应性与鲁棒性。如表 6 所示，所提方法在 PlantSeg 上 mIoU 达 77.99%，相较 DeepLabV3Plus 提升 4.3 个百分点，体现出良好的跨植物病斑类型的泛化能力。可视化结果如图 7 所示，图像 1 中，除 SeaFormer－B 和本文模型外，其余模型均对左侧小框内区域存在误分，且 AISOA－SSFormer 与 DeepLabV3Plus 在右侧小框内的病斑分割不够精细；图像 2 中，本文模型分割精

表 6 不同模型在 PlantSeg 数据集上测试结果

Tab. 6 Testing results of different models on PlantSeg dataset

模型	平均交并比/%	准确率/%	参数量
DeepLabV3Plus(ResNet－50)	73.69	84.25	2.679×10^7
PSPNet (ResNet101)	70.45	82.68	6.700×10^7
SegNext－T	76.06	85.94	4.300×10^6
SeaFormer－B	75.23	85.11	8.600×10^6
SCTNet－B	76.82	86.22	1.740×10^7
本文模型	77.99	88.21	2.512×10^7

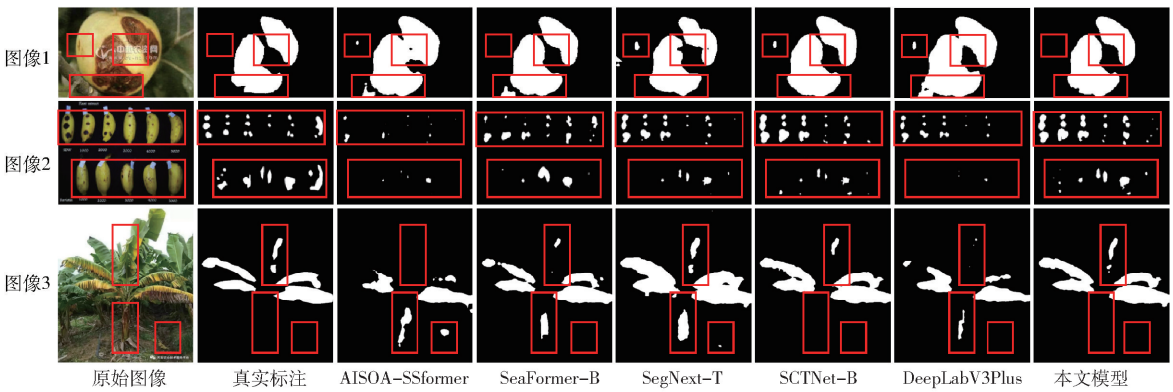


图 7 PlantSeg 数据集可视化效果

Fig. 7 PlantSeg dataset visualization

度更高，检测到的病斑区域更多；图像 3 中，除 SCTNet－B 和本文模型外，其余模型均将树干误判为病斑，且存在分割不完整的问题。上述结果进一步验证了本文模型的有效性与鲁棒性。

3　结论

（1）提出了层次特征与语义融合解码的苹果叶片病害语义分割模型，以应对微小病斑检测、尺度多样性及边界模糊等挑战。

（2）模型设计中，HFN 骨干网络结合 CNN 与 Transformer 结构，逐层提取病斑特征，有效捕捉微小病斑及全局语义信息；ES－ASPP 模块融合空洞卷积与边缘引导机制，增强尺度感知与结构表达能力；SFD 解码器高效融合不同层级语义特征，实现高效边缘还原与准确病斑分割。

（3）所提方法在自建数据集和 PlantSeg 公开数据集上 mIoU 分别达到 91.13% 与 77.99%，验证了模型在病害分割任务中的有效性。

参　考　文　献

[1] Zhao Y, Li X, Wang J, et al. Metabolic and cardiovascular benefits of apple and apple-derived products: a systematic review and meta-analysis[J]. Front. Nutr., 2022, 9: 766155.

[2] Jia J, Zhang Y, Wang H, et al. RAAWC－UNet: an apple leaf and disease segmentation method based on residual attention and atrous spatial pyramid pooling improved U－Net with weight compression loss[J]. Front. Plant Sci., 2024, 15: 1305358.

[3] Storey G, Meng Q, Li B. Leaf disease segmentation and detection in apple orchards for precise smart spraying in sustainable agriculture[J]. Sustainability, 2022, 14(3): 1458.

[4] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770－778.

[5] Howard A G, Sandler M, Chu G, et al. Searching for MobileNetV3[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision(ICCV), 2019: 1314－1324.

[6] He K, Gkioxari G, Dollar P, et al. Mask R－CNN[C]// Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017: 2961－2969.

[7] Agarwal M, Gupta S K, Biswas K K. Plant leaf disease segmentation using compressed UNet architecture[C]// Trends and Applications in Knowledge Discovery and Data Mining, 2021: 9－14.

[8] Ronneberger O, Fischer P, Brox T. U－Net: convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer-Assisted Intervention(MICCAI), 2015: 234－241.

[9] Li L, Wang B, Li Y, et al. Diagnosis and mobile application of apple leaf disease degree based on a small-sample dataset[J]. Plants, 2023, 12(4): 786.

[10] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]// Proceedings of the European Conference on Computer Vision(ECCV), 2018: 801－818.

[11] 鲍文霞，赵诗意，黄林生，等. 基于轻量级密集多尺度注意力网络的小麦叶部锈病识别方法[J]. 农业机械学报，2024，55(11)：21－31.
Bao Wenxia, Zhao Shiyi, Huang Linsheng, et al. Lightweight dense multi-scale attention network for identification of rust on wheat leaves[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 21－31. (in Chinese)

[12] Zeng T W, Li C M, Zhang B, et al. Rubber leaf disease recognition based on improved deep convolutional neural networks with a cross-scale attention mechanism[J]. Frontiers in Plant Science, 2022, 13: 829479.

[13] 邹一波，周泽政，陈明，等. 基于注意力机制的植物三维点云语义分割方法[J]. 农业机械学报，2025，56(3)：129－139，157.
Zou Yibo, Zhou Zezheng, Chen Ming, et al. 3D plant point cloud semantic segmentation method APSegNet based on attention mechanism[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 129－139, 157. (in Chinese)

[14] Ding Y, Yang W, Zhang J. An improved DeepLabV3 + based approach for disease spot segmentation on apple leaves[J]. Computers and Electronics in Agriculture, 2025, 231: 110041.

[15] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]// Advances in Neural Information Processing Systems (NeurIPS), 2017: 5998－6008.

[16] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 7132－7141.

[17] Larsson G, Maire M, Shakhanrovich G. FractalNet: ultra-deep neural networks without residuals[C]// Proceedings of the 5th International Conference on Learning Representations, 2017.

[18] Zhang X, Zhou X, Lin M, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 6848－6856.

[19] Wei T, Chen Z, Yu X, et al. PlantSeg: a large-scale in-the-wild dataset for plant disease segmentation[PP/OL]. V1. arXiv(2024－09－06)[2025－02－06]. https://doi.org/abs/2409.04038.

[20] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 2881－2890.

[21] Guo M H, Lu C Z, Hou Q, et al. SegNeXt: rethinking convolutional attention design for semantic segmentation[J]. Advances in Neural Information Processing Systems(NeurIPS), 2022, 35: 1140－1156.

[22] Wan Q, Huang Z, LU J, et al. SeaFormer: squeeze-enhanced axial transformer for mobile semantic segmentation[C]// Proceedings of the International Conference on Learning Representations (ICLR), 2023.

[23] Dai W, Zhou W, Guo Z, et al. AISOA－SSformer: an effective image segmentation method for rice leaf disease based on the transformer architecture[J]. Plant Phenomics, 2024, 6.

[24] Xu Z, Wu D, Yu C, et al. SCTNet: single-branch CNN with transformer semantic information for real-time segmentation[C]// Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence(AAAI), 2024: 6378－6386.