

DOI:10.19651/j.cnki.emt.2210583

结合空频域多尺度特征的植株叶片分割方法^{*}

罗薇^{1,2} 陈峰³ 张晓伟³ 梁虹^{1,2} 李军营³

(1. 云南大学信息学院 昆明 650504; 2. 云南省高校物联网技术及应用重点实验室 昆明 650504;
3. 云南省烟草农业科学研究院 昆明 650021)

摘要: 为实现植物表型参数的精准获取,针对自然环境下不同尺度叶片分割的细节损失问题,提出结合空频域多尺度特征的植株叶片分割方法。以 U-Net 网络架构为基础,设计下采样频域变换模块,在卷积神经网络中引入频域特征表示替换池化层,利用 2D-DCT 和 2D-IDCT 的频域变换方法感知植株叶片目标的全局语义特征;构建多尺度特征融合模块,增加 6 个上采样节点,提取和连接植株叶片图像细粒度的特征信息;改进通道注意力模块学习分支特征,采用联合损失函数优化网络性能。在 2017CVPPP 公开数据集上开展实验,结果表明,植株叶片分割网络的交叉比、平均交叉比、像素准确率、精准率和 F-score 分数分别达到了 97.07%、98.04%、99.53%、99.68% 和 99.74%。与 FCN-8s、FCN-ResNet、DeepLabV3+、SegNet 和 U-Net 模型相比,网络的交叉比和平均交叉比最高提升 23.32% 和 12.43%,在较小尺度和细节处理上改善了植株叶片的分割精度,可为植物表型方向的应用研究提供一种可借鉴的思路。

关键词: 植物表型;植株叶片分割;U-Net;频域特征表示;多尺度特征

中图分类号: TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.6

Plant segmentation method based on multi-scale fusion network of spatial-frequency domain features

Luo Wei^{1,2} Chen Feng³ Zhang Xiaowei³ Liang Hong^{1,2} Li Junying³

(1. School of Information, Yunnan University, Kunming 650504, China;

2. University Key Laboratory of Internet of Things Technology and Application, Yunnan Province, Kunming 650504, China;

3. Yunnan Academy of Tobacco Agriculture Science, Kunming 650021, China)

Abstract: To improve the effect of plant segmentation and achieve accurate acquisition of plant phenotypic parameters, this study proposes a plant segmentation network to fuse spatial-frequency domain feature representation. Based on the U-Net network architecture, the frequency domain transform module is built for down-sampling, the frequency domain feature representation is introduced in the convolutional neural network to replace the pooling layer, and the frequency domain transform method of 2D-DCT and 2D-IDCT is used to perceive the global semantic features of the plant. The multi-scale feature fusion module is constructed by adding six up-sampling nodes to extract and connect the fine-grained feature information of the plant image. The channel attention module is modified to learn branch features and the hybrid loss function is employed to optimize the network. Experiments are conducted on the 2017 CVPPP public dataset, and the results show that the intersection over union, mean intersection over union, pixel accuracy, precision and F1 of the plant segmentation network reach 97.07%, 98.04%, 99.53%, 99.68% and 99.74%, respectively. Compared with the FCN-8s, FCN-ResNet, DeepLabV3+, SegNet and U-Net models, the intersection over union and mean intersection over union of the network were improved by up to 23.32% and 12.43%. The proposed method can improve the segmentation accuracy of plant at smaller scales in detail processing, and it can provide a useful idea for applied research in the plant phenotype.

Keywords: plant phenotype; plant segmentation; U-Net; frequency domain feature representation; multi-scale features

0 引言

植物表型是指植物在受基因、管理和环境因素相互作

用后,表现出来结构及生长的特征和性状,植物表型参数的测量及分析可以有效调节植物的育种环节,最终影响植物的生长发育和作物产量。植物表型参数获取工作主要通过

收稿日期:2022-07-06

^{*} 基金项目:云南省烟草公司重大科技项目(2021530000241025)、云南大学第十三届研究生科研创新项目(2021Y259)资助

手工测量结合软件分析来完成,这些方法往往耗时费力并且结果准确性较低^[1]。近年来,基于计算机视觉技术的植物表型研究方法广泛用于植物生长监测、叶片分割及计数及病虫害分类识别等领域,实现在复杂环境下植株和植株叶片的精确分割是在各研究方法中最为基础和关键的一步^[2]。

现阶段大部分植株叶片分割方法的研究人员使用在受控实验室和温室环境下采集的图像数据,主要代表为拟南芥和烟草花两种植物类型^[3-6]。传统植株叶片分割方法通常借助目标叶片的颜色纹理特征,通过阈值分割等算法实现植株叶片分割^[7]。但当面对图像数据不同的保真度、分辨率及场景复杂性等挑战时,因缺乏深层语义特征的表征能力,传统图像处理方法在性能提升上存在一定的局限性。随着深度学习技术的发展,具备强大特征学习能力的深度学习模型现已开始应用于植株叶片分割领域。Kuznichov 等^[8]提出一种尽可能保留植株几何特征的数据增强方法,在基于 Mask R-CNN 模型的实验中进一步提高了叶片分割的准确性,但该方法需要在一定的数据约束环境下进行。Aich 等^[9]采用深度反卷积和卷积网络相结合的方法实现了对莲座形植株叶片的分割,该方法采用神经网络进行特征提取并直接用于像素级别的分类,忽略了浅层特征对于细节预测的贡献程度。廖娟等^[10]提出一种利用特征金字塔融合目标区域语义和边缘信息的分割模型,该方法在对不同的作物苗期植株分割时表现出良好的泛化性,但它引入边缘监督及相应的冲突弱化结构却增加了网络模型的计算复杂度。如何在提升效率的同时,最大限度利用多尺度特征信息来减少植株叶片分割时的细节损失是一个重要的研究方向。

基于深度学习的方法在图像分割技术研究上表现出明显的优势,这些网络模型所提取的浅层特征包含更多的细节信息但受噪声影响较大,深层特征具有更强的语义信息但对细节的感知能力较差,融合不同尺度的特征是提升分割性能的一个重要手段^[11-12]。其中,将具体图像的频域表示引入卷积神经网络(convolutional neural network, CNN)在特征提取方面已展现出巨大的潜力,此时的图像特征图可设定为任意的输出维度,并且能够适当控制维度的缩减。Adam 等^[13]采用了一种全新的频带限制训练方法,可以有效优化算力资源,同时保留原始网络架构和高精度。Pratt 等^[14]提出了傅里叶卷积神经网络(Fourier convolution neural network, FCNN),该网络在可行的计算时间内类比空间域采用不同的操作方法,在傅里叶域中训练并取得了令人满意的性能。Ayat 等^[15]提出了一种用于激活函数的新型谱校正线性单元(spectral rectified linear unit, SReLU),避免了多个计算密集型域变换,进一步优化网络的分类速度。Pour 等^[16]提出通过将拉普拉斯金字塔注入卷积神经网络来整合变换域表示,节约训练时间的同时还进一步提升了性能。考虑到植物表型参数获取的实际工作

场景,植株叶片分割网络不宜过大过深,不仅要利用图像中目标叶片的多尺度特征信息,还要保证一定的分割速度。

针对植株叶片分割中目标叶片位置、边缘等细节信息的损失问题,本文提出一种结合空频域多尺度特征的植株叶片分割方法,结合空间域和频域的语义表征能力,将频域表示引入卷积神经网络当中,实现植株叶片与背景的精确分割。以 U-Net 网络模型结构^[17]为基础,利用频域变换和截断操作舍弃高频系数,保留频率分量中的低频部分,在编码器的子网络中融合下采样频域变换模块,保留更多植株叶片性状信息的同时弱化多次下采样带来的细节损失;基于跳跃连接结构,适当增加上采样节点,引导网络在解码器的子网络中聚合多尺度语义特征,增强网络对植株叶片各尺寸对象的关注;增加轻量级通道注意力模块,在多解码通道的依赖关系中学习更多的重点特征内容,并采用联合损失函数优化和提升网络整体性能,最终解决植株叶片分割精度不高的问题,为植株叶片分割研究提供一种新的思路。

总体来看,本文的创新及贡献主要包括以下两个方面:在现有的植株叶片分割方法中,首次将频域特征表示方法引入改进后的多尺度语义分割网络当中,实现下采样频域变换模块与原模型架构的有效融合。提出一种结合空频域多尺度特征的植株叶片分割方法,加入跳跃连接和注意力机制,采用联合损失函数加以优化,在公开图像数据集上验证了该方法在较小尺度和细节处理方面具有良好的分割效果。

1 材料与方法

1.1 数据集来源

本文采用公开的 2017CVPPP (computer vision problems in plant phenotyping)数据集^[4]。采用将实验植株置于花盆或托盘中并从顶部拍摄的方法,图像均由 Canon AD1000 Power-Shot 设备采集,主要包含拟南芥和烟草花植株品种且不同尺寸的 5 个(A1~A5)数据集。其中,A1 和 A2 分别为野生的拟南芥植株图像和突变后叶片形状各异的拟南芥植株图像;A3 为烟草花植株图像;A4 为拟南芥植株生长发育至成熟的图像;A5 为前 4 个子数据集的总和。实验选取 A5 数据集共 810 张,具体介绍参如表 1 所示。除植株叶片对象自身的多样性,数据集还建立在不同的环境复杂条件下,如存在水面引起的反射、背景和异类物体的变化及叶片的遮挡情况等,在植株叶片分割任务中存在一定的挑战性。A1~A4 子数据集的部分图像如图 1 所示。

表 1 2017CVPPP 数据集属性

数据集	植株品种	图像尺寸	图像数量
A1	拟南芥	500×530	128
A2	拟南芥	530×565	31
A3	烟草花	2 448×2 048	27
A4	拟南芥	441×441	624

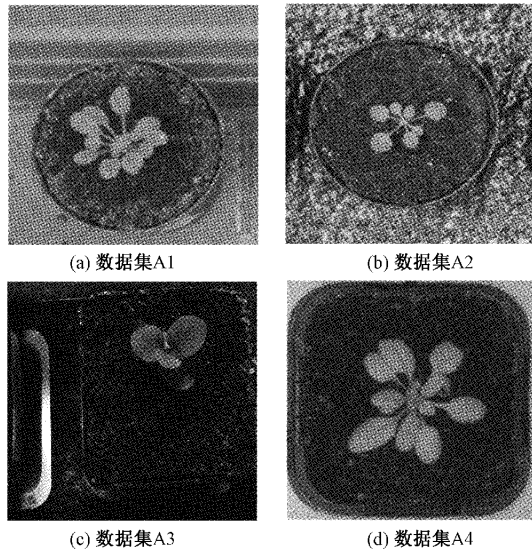


图 1 2017CVPPP 数据集部分图像

1.2 植株叶片分割网络结构

由于植株叶片图像中存在大小和形状各异的叶片目标,为保留目标群体更多的细节信息,提高植株叶片分割网络提取多尺度特征的能力,本研究基于 U-Net 模型的编解码架构提出改进方法,搭建一个融合空频域多尺度特征的植株叶片分割网络,网络的整体结构如图 2 所示。网络由下采样频域变换模块、多尺度特征融合模块和通道注意力模块 3 部分构成,其中,下采样频域变换模块将具体变换依次嵌入编码过程,提升编码器对全局植株及叶片细节特征提取的能力;多尺度特征融合模块主要是由密集连接的上采样节点所组成的解码子网络结构,用于提取并聚合植株叶片的多尺度语义特征;通道注意力模块主要用于学习和关注重要特征信息,优化和提升植株叶片分割网络的精度。

1) 下采样频域变换模块

近期有研究表明,压缩的频域特征表示能够有效抑制空间域当中的噪声,并且丰富图像的理解模式^[18]。在标准

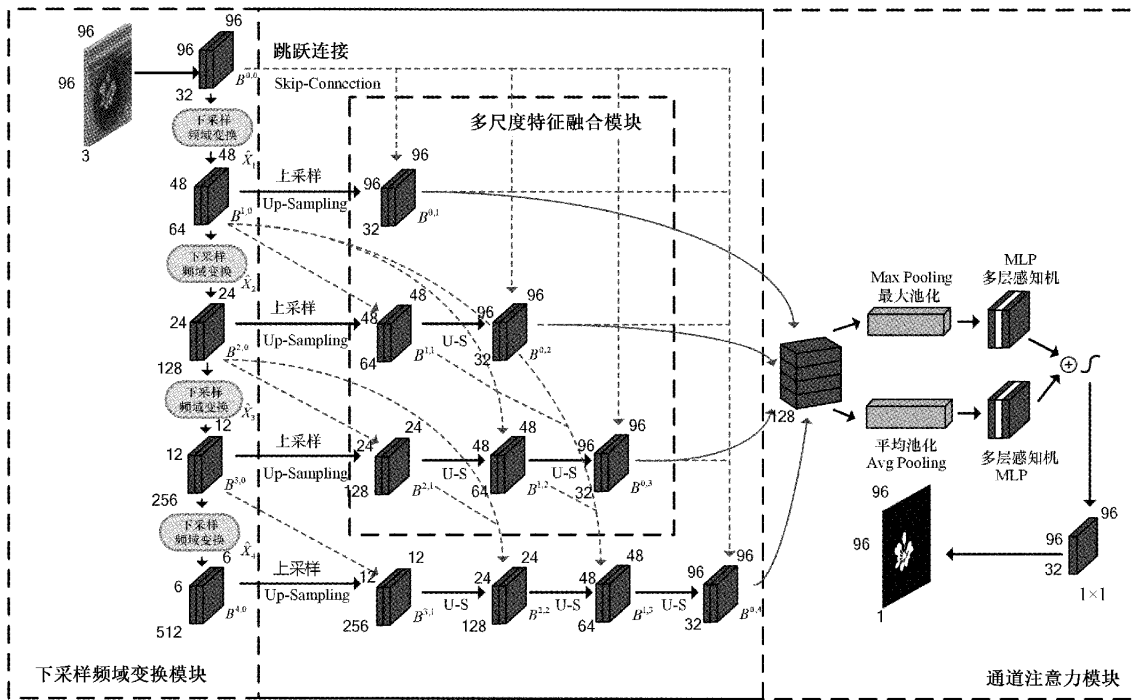


图 2 植株叶片分割网络结构

的卷积操作中,所有输入和输出特征图都局限在空间域,其中每个特征像素仅代表其自身或者很小的局部区域。相反,由于频域变换有其明确的物理意义,当中每个像素都包含整个特征图的全局信息,图像可以在快速傅里叶变换(fast Fourier transform, FFT)、离散余弦变换(discrete cosine transform, DCT)、小波(wavelet)及其他域中进行表示。在数字图像处理领域,二维离散余弦变换(2D-discrete cosine transform, 2D-DCT)具备去相关性和能量集中的特性,是一种解决信号和图像数据压缩问题的有效途径。输入图像 $x_{hw} \in \mathbb{R}^{H \times W}$ 和

二维离散余弦变换 $F_{pq}(x) \in \mathbb{C}^{H \times W}$ 的定义为:

$$F_{pq}(x) = \alpha_p \alpha_q \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} x_{hw} \cos \frac{\pi(2h+1)p}{2H} \cos \frac{\pi(2w+1)q}{2W}, \quad \begin{matrix} 0 \leq p \leq H-1 \\ 0 \leq q \leq W-1 \end{matrix} \quad (1)$$

$$x_{hw} = \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} \alpha_p \alpha_q F_{pq}(x) \cos \frac{\pi(2h+1)p}{2H} \cos \frac{\pi(2w+1)q}{2W}, \quad \begin{matrix} 0 \leq h \leq H-1 \\ 0 \leq w \leq W-1 \end{matrix} \quad (2)$$

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{H}}, & p = 0 \\ \sqrt{\frac{2}{H}}, & 1 \leq p \leq H-1 \end{cases} \quad (3)$$

$$\alpha_q = \begin{cases} \frac{1}{\sqrt{W}}, & q = 0 \\ \sqrt{\frac{2}{W}}, & 1 \leq q \leq W-1 \end{cases} \quad (4)$$

其中, α_p 和 α_q 是系数矩阵, H 和 W 分别为输入图像像素的高度和宽度。

卷积神经网络中空间域池化方法的主要功能包括降低维度、消除冗余和压缩特征,本文参考 Xu 等^[19]将离散余弦变换用于在频域中提取特征的方法,对二维离散余弦变换域的特征表示图进行截断,选取直流(direct current, DC)系数,保留图像低频信息,以此在变换域中实现类似池化的操作。即给定输入特征张量 $X \in \mathbb{R}^{H_s \times W_s \times C_s}$, H_s 、 W_s 和 C_s 分别表示该张量在空间域的维度大小及特征通道数,期望得到一个维度减半的输出映射 $\hat{X} \in \mathbb{R}^{\frac{H_f}{2} \times \frac{W_f}{2} \times C_f}$, 此时 H_f 、 W_f 和 C_f 分别表示该张量在频域的维度大小及特征通道数。下采样频域变换相当于一个低通滤波器,为特征提取提供了更广阔的阈值空间,而空间域池化后的特征图大部分功率谱都处在较低的频率上,不利于最大化地学习特征。图 3 对比了将下采样频域变换和空间域池化方法(最大池化和平均池化)应用于单通道图像的不同结果,可以看出前者突出强调了植株叶片更为准确的形状、边缘等细节信息,加大了前景与背景区域的区分度,后者基本保留了植株叶片的浅层特征信息。

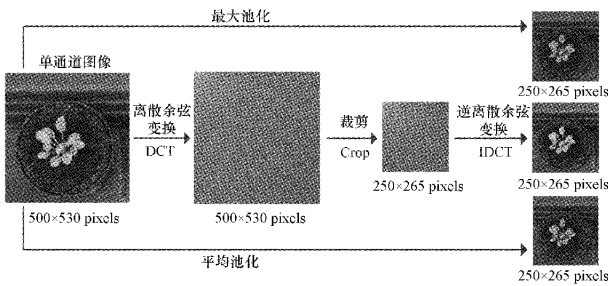


图 3 单通道图像的下采样方法对比

下采样频域变换算法的实现步骤十分简洁,如算法 1 所示,当输入的空间域特征图进行二维离散余弦变换以后,得到变换后的频域特征图,然后每个频率维度的特征图进行裁剪,最终进行二维离散余弦逆变换得到期望的特征图。在图 2 的下采样频域变换模块中,针对原 U-Net 模型的编码部分,卷积块 $B^{i,j}$ ($i = 0, 1, 2, 3; j = 0$) 涉及到 2×2 的最大池化(Max Pooling)下采样操作均改进为下采样频域变换算法,进而完整构建得到网络模型的下采样频域变换模块。

算法 1 下采样频域变换

输入 $X \in \mathbb{R}^{H_s \times W_s \times C_s}$
 输出 $\hat{X} \in \mathbb{R}^{\frac{H_f}{2} \times \frac{W_f}{2} \times C_f}$
 1. $X_{in} \leftarrow F(X)$
 2. $X_{out} \leftarrow Crop(X_{in}, \frac{H}{2} \times \frac{W}{2} \times C)$
 3. $\hat{X} \leftarrow F^{-1}(X_{out})$

2) 多尺度特征融合模块

Zhou 等^[20]提出的 UNet++ 通过跳跃连接将浅层的细粒度空间特征传输到子解码器,有效提升了模型的精度,这意味着在深层应用浅层的特征信息是有效的。针对植株叶片目标尺度不均进而影响分割性能的问题,本研究将下采样频域变换模块输出不同尺度的 4 个特征张量 $b^{i,j}$ ($i = 1, 2, 3, 4; j = 0$) 分别输入同级别的上采样节点进行恢复,增加 6 个上采样节点,利用高度灵活的跳跃连接结构和不同深度共享解码器的高效集成,实现不同语义尺度特征的高度融合。

如图 2 所示,在多尺度特征融合模块中,卷积块通过上采样的子解码器将特征图恢复到原始大小。此外,浅层子解码器中模块的输出先与相同大小的更深层子解码器中的模块结合后再开始下一步的上采样操作。其中,卷积块 $B^{i,j}$ 主要包含两次 3×3 的卷积操作、批量归一化(batch normalization, BN)操作和 ReLU 激活函数运算^[21]。其定义如下:

$$y_1 = \sigma(F(x, \{W_i\})) \quad (5)$$

$$y = \sigma(F(y_1, \{W_i\})) \quad (6)$$

其中, x 和 y 是卷积块的输入和输出张量,函数 $F(x, \{W_i\})$ 表示待训练的堆叠非线性层,此处为简化符号省略了偏差值。最终对输出张量增加一个非线性层运算, $\sigma(\cdot)$ 为 ReLU 激活函数。

$b^{i,j}$ 表示卷积块 $B^{i,j}$ 输出的特征图映射,其中 i 表示输入多尺度特征融合模块的级数, j 表示沿跳跃连接卷积块的上采样级数。 $b^{i,j}$ 用公式表示为:

$$b^{i,j} = \begin{cases} \mathcal{H}(\mathcal{F}(b^{i-1,j})), & j = 0 \\ \mathcal{H}([\![b^{i,k}]_{k=0}^{i-1}\!]^{-1}, \mathcal{U}(b^{i+1,j-1})), & j > 0 \end{cases} \quad (7)$$

其中, $\mathcal{H}(\cdot)$ 表示卷积块操作,此时将输入特征图的初始尺寸大小为 $96 \text{ pixels} \times 96 \text{ pixels} \times 32$ 通道。 $\mathcal{F}(\cdot)$ 通过频域变换实现了下采样的功能。 $\mathcal{U}(\cdot)$ 表示由双线性插值方法实现的上采样操作。 $[\![\cdot]\!]^{-1}$ 表示将两个特征图在第 2 个维度上进行通道的连接。在图 2 中, $j = 0$ 级的卷积块向下采样和提取特征, $j > 0$ 级的卷积块接收 $j + 1^{\text{th}}$ 的输入,其中 j^{th} 的输入都是来自相同跳跃连接的前一个卷积块的输出, $j + 1^{\text{th}}$ 的输入是下跳连接的上采样输出。

3) 通道注意力模块

针对植株叶片分割任务中受无关特征信息影响较大

进而损失较多边缘特征信息的情况,本研究加入轻量级通道注意力模块,对各通道的特征进行权重加权,强调和关注植株叶片的边缘特征信息,改善叶片边缘的分割效果。为了对多尺度特征融合模块输出的分支特征图进行优化,受 Woo 等^[22]提出融合通道和空间注意力的注意力模块(convolutional block attention module, CBAM)的启发,如图 2 所示,本研究将 4 个分支特征图 $b^{i,j}$ ($i=0; j=1,2,3,4$) 在第 2 个维度上进行拼接操作,融合为一个 $96 \text{ pixels} \times 96 \text{ pixels} \times 128$ 通道大小的特征图,然后将其输入到改进后的轻量级通道注意力模块(channel attention module-4, CAM-4)对分支间的权重关系进行全局有效的感知,获取得到的特征图尺寸为 $96 \text{ pixels} \times 96 \text{ pixels} \times 32$ 通道,最终经过 1×1 的卷积运算输出分割结果预测图。注意力模块 CAM-4 在充分提取植株叶片特征的同时调整空间特征维度,增大全局感受视野,具体的运算过程如下:

$$M_{CAM-4}(f) = \sigma(W_1(W_0(f_{avg})) + W_1(W_0(f_{max}))) \quad (8)$$

其中, $\sigma(\cdot)$ 表示 sigmoid 激活函数, f_{avg} 和 f_{max} 分别表示平均池化(Mean Pooling)和最大池化(Max Pooling),平均池化实现输入特征图空间维度的压缩和空间信息的聚合,最大池化用于提取独特的物体特征来推断通道上的注意力。该运算先将两种空间域池化方法计算得到的特征分别输入到多层感知(multilayer perception, MLP)组成的共享网络当中,然后将多层感知计算的结果输入 sigmoid 激活函数得到通道注意力图 M_{CAM-4} 。此处的多层感知网络由 2 个全卷积层(fully convolutional networks, FCN)和 ReLU 激活函数构成, $W_0 \in \mathbb{R}^{C/r \times C}$ 和 $W_1 \in \mathbb{R}^{C \times C/r}$ 分别表示权重, C 和 r 分别为特征图的通道数和通道参数。

1.3 联合损失函数

考虑到植株叶片图像中,目标植株及叶片所占区域像素较少,而背景所占区域像素较多。为应对语义分割中正负样本失衡的场景,本文训练网络模型过程中采用联合损失函数进行优化,损失函数定义为 Dice 损失函数与二分类交叉熵损失函数(binary cross entropy, BCE)的联合损失,前者从全局进行均衡,后者在像素级别上计算,表达式为:

$$l_{loss} = l_{dice} + \frac{1}{2} \times l_{bce} \quad (9)$$

Dice 损失函数 l_{dice} 广泛应用于医学影像分割领域^[23],当前景像素与背景像素的占比差距较大时取得的结果较好,定义如下:

$$l_{dice} = 1 - \frac{2 \times (\hat{y} \cap y)}{\hat{y} \cup y} \quad (10)$$

式中: \hat{y} 为分割预测图的像素点集合, y 为分割目标图的像素点集合, \cap 为两个像素集合的交集运算, \cup 为两个像素集合的并集运算。当出现同一植株图像分割叶片内容不均衡时,通过将其与二分类交叉熵损失函数进行加权融

合以后,可以有效解决小目标被忽略和无法学习到正确梯度下降方向的问题。二分类交叉熵损失函数 l_{bce} 计算如下:

$$l_{bce} = -\frac{1}{N} \sum_{i,j} (y_{i,j} \ln(\hat{y}_{i,j}) + (1 - y_{i,j}) \ln(1 - \hat{y}_{i,j})) \quad (11)$$

式中: N 为训练一个批量中像素数的总和, $y_{i,j} \in \{0,1\}$ 表示分割目标图中坐标为 (i,j) 的像素点, $\hat{y}_{i,j} \in \{0,1\}$ 表示分割预测图中坐标为 (i,j) 的像素点,植株叶片分割图中前景像素的值标注为 1,背景像素的值标注为 0。

2 实验及结果分析

2.1 实验环境

实验选用 Intel (R) Core (TM) i9-10900F CPU @ 2.80 GHz 2.81 GHz, GPU 配备 NVIDIA GeForce RTX 3080,显存 10 GB。软件环境为 Windows 64 位操作系统, CUDA 版本为 11.4.136,基于 Pytorch 深度学习框架使用 Python 编程语言实现。

2.2 模型评价指标

本文采用评价分割模型性能的 5 个常用指标,包括交叉比(intersection over union, IoU)、平均交叉比(mean intersection over union, MIoU)、像素准确率(pixel accuracy, PA)、精准率(precision, P)和 F-score 分数(F1)。在该研究中,交叉比 IoU 是指模型对植株叶片分割结果和真实结果的像素交集与像素并集之比,平均交叉比 MIoU 是指植株叶片像素交并比和背景像素交并比的平均值,像素准确率 PA 表示模型正确预测的像素数占总像素数的比例,精准率 P 是指模型正确分割的像素数占所有预测为该分割类别像素数的比例, F-score 分数综合考虑了植株叶片分割的准确率和召回率,是衡量二分类模型性能的一个重要指标。各评价指标计算公式如下:

$$IoU = \frac{T_P}{F_P + T_P + F_N} \quad (12)$$

$$MIoU = \frac{1}{k+1} \cdot \frac{T_P}{F_P + T_P + F_N} \quad (13)$$

$$PA = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (14)$$

$$P = \frac{T_P}{F_P + T_P} \quad (15)$$

$$F_1 = 2 \times \frac{P \cdot recall}{P + recall} \quad (16)$$

2.3 实验参数设置及训练过程分析

为排除可能存在的干扰因素,本文所有实验均在相同的参数配置下进行,采用 2017CVPPP 数据集,训练集和测试集的数量按 8:2 比例进行划分,并使用随机旋转 90° 、垂直或水平翻转、添加高斯噪声及随机更改图像值等数据增强方法对原始图像数据进行扩充。在训练网络时将初

始学习率设为 0.000 1,使用余弦退火学习率(cosine annealing LR)衰减策略,批次大小设置为 16,训练轮次为 100 次,采用 Adam 算法对网络进行优化。

本研究提出的网络模型在训练过程中测试得到的交叉比和损失值变化曲线如图 4 所示,随着训练轮次的增加,IoU 值逐渐提升,损失值逐渐降低,观测指标在前 10 次迭代过程中变化较明显。当训练轮次到达 100 次时,模型基本收敛且取得较好的训练效果,交叉比 IoU 的值较为稳定。

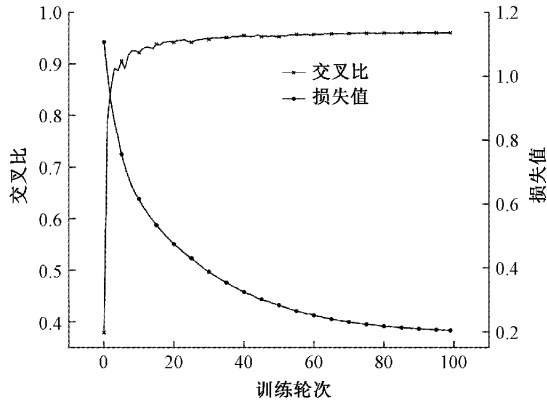


图 4 网络训练中交叉比和损失值变化曲线

2.4 实验结果分析

1)网络模型对比分析

为验证本文方法在植株叶片分割任务中优于常用的深度学习方法,在相同的实验参数设置下分别对本文模型与 FCN-8s^[24]、FCN-ResNet^[24]、DeepLabV3+^[25]、SegNet^[26]、U-Net^[17]主流模型开展训练和测试,表 2 为 7 种分割模型各项评价指标的对比结果。

表 2 在不同模型上植株叶片的分割结果 %

分割模型	交叉比 IoU	平均 交叉比 MIoU	像素 准确率 PA	精准率 P	F 分数 F ₁
FCN-8s	77.38	87.23	96.27	96.27	97.47
FCN-ResNet	85.26	91.76	97.35	97.79	98.37
DeepLabV3+	87.79	93.18	97.85	98.18	98.66
SegNet	92.09	95.10	98.45	99.09	99.04
U-Net	94.34	96.56	98.92	99.15	99.33
本文模型	97.07	98.04	99.53	99.68	99.74

由表 2 可以看出,本文提出的基于空频域特征多尺度融合网络的植株叶片分割方法性能最优,各项评价指标较 FCN-8s、FCN-ResNet、DeepLabV3+、SegNet、U-Net 模型均有提升。其中,本文网络的 IoU、MIoU、PA、P 和 F1 值分别比 U-Net 模型高出了 2.73%、1.48%、0.61%、0.53% 和 0.41%。

本文网络在 IoU 及 MIoU 评价指标上分别达到 97.07% 和 98.04%,与其他模型相比最高提升 19.69% 和 10.81%。可见,本文设计频域变换方法引入改进后的多尺度语义分割网络,有利于网络在训练过程中提取和学习细节特征,以此改善植株叶片分割模型的性能。

图 5 为 6 种不同分割模型的植株叶片分割效果,本文模型和其他分割模型都完成了植株叶片分割任务。FCN-ResNet 和 DeepLabV3+ 网络整体分割效果略优于 FCN-8s,分割出来的植株叶片边界相较而言也更为清晰,但都基本没有定位到植株的叶柄部位。SegNet 和 U-Net 网络针对目标叶片的分割效果较好,但在图像 4 和 5 中对植株

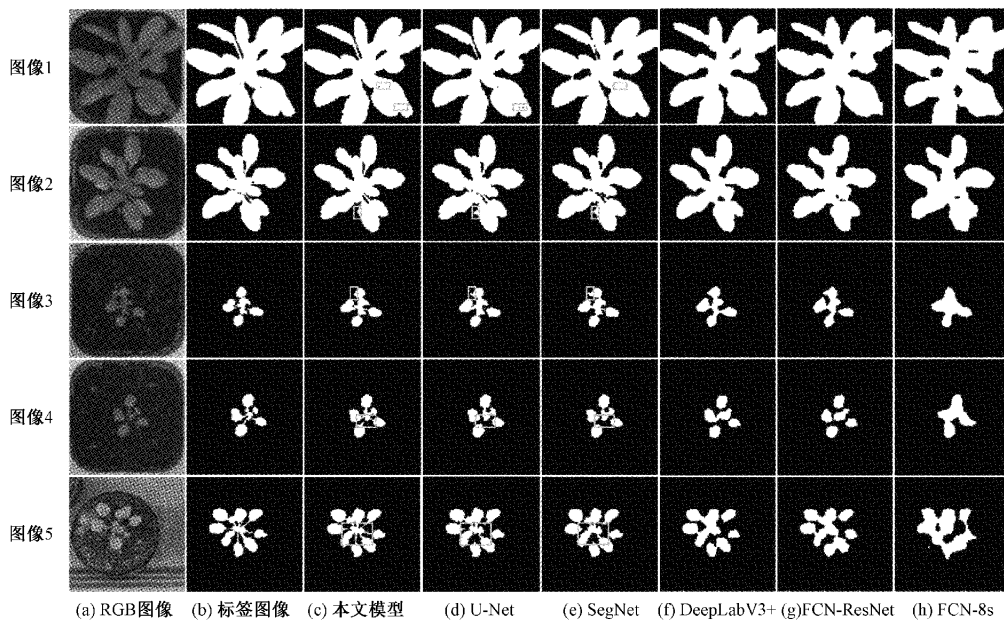


图 5 网络模型的植株叶片分割效果比较

叶柄部位的分割效果较差;在图像 2 和 3 中明显受到土壤异类物体的干扰使得植株叶片边界分割模糊;在图像 1 中叶片内部的极少像素点均出现了未正确分割的情况。总体来看,本文网络对植株叶片的分割效果最好,尤其是在小尺度分割和细节处理方面,均优于其他网络结构。

2) 频域变换对模型性能的影响

为验证本文设计的植株叶片分割模型中引入频域变换方法的有效性,针对测试集的样本数据在下采样频域变换模块中分别采用最大池化和频域变换结构对模型进行性能测试,2 种下采样结构测试得到的模型性能参数结果如表 3 所示。

表 3 下采样结构微调实验 %

下采样结构	交叉比 (IoU)	平均交叉比 (MIoU)	像素准确率 (PA)	精准率 (P)	F 分数 (F_1)
最大池化	95.06	97.21	99.13	99.29	99.46
频域变换	97.07	98.04	99.53	99.68	99.74

表 3 中,网络的下采样结构为频域变换时取得的各项评价指标均优于采用最大池化方法,其中频域变换结构在测试中获得的 IoU 指标相对于最大池化结构提升了 2.01%,表明将频域特征表示方法应用于多尺度语义分割网络进行训练,能够增加网络对植株叶片性状信息的感知,抑制深层特征的细节损失,有效改善语义分割的最终效果。

3) 注意力模块对模型性能的影响

为验证本文改进后的通道注意力模块 CAM-4 优于其他轻量级的注意力模块,在模型通道注意力模块部分的不同位置分别嵌入注意力 SE(squeeze-and-excitation)^[27]和 S2-MLPv2(spatial-shift-multilayer perceptronv2)^[28],进行对照实验,实验结果如表 4 所示。

表 4 注意力模块微调实验 %

注意力模块	交叉比 (IoU)	平均交叉比 (MIoU)	像素准确率 (PA)	精准率 (P)	F 分数 (F_1)
S2-MLPv2	94.85	96.82	99.00	99.27	99.38
SE	96.69	97.89	99.49	99.69	99.72
CAM-4	97.07	98.04	99.53	99.68	99.74

由表 4 可知,通过将模型的通道注意力模块部分设置为通道注意力模块 CAM-4 后,模型的各项性能指标都有不同程度的提升。其中,与 S2-MLPv2 相比,IoU 和 MIoU 指标分别提升 2.22% 和 1.22%,表明通道注意力模块 CAM-4 有利于各解码分支通道间重点特征的学习。

4) 联合损失函数对模型性能的影响

本文提出的网络模型在训练过程中,损失函数会依据分割预测结果产生的误差进行反向传播引导各项参数进行学习 and 优化。为验证本文设计的联合损失函数对模型的优化效果,在网络的训练过程中分别设置二分类交叉熵损失函数 BCE、Dice 损失函数及联合损失函数对模型开展性能测试,测试结果如表 5 所示。

表 5 损失函数微调实验 %

损失函数	交叉比 (IoU)	平均交叉比 (MIoU)	像素准确率 (PA)	精准率 (P)	F 分数 (F_1)
BCE	95.10	97.25	99.20	99.42	99.48
Dice	94.98	97.14	99.15	99.26	99.43
BCE-Dice	97.07	98.04	99.53	99.68	99.74

表 5 中,对比本文网络模型经过 3 种损失函数训练后的分割性能,联合损失函数取得的 IoU 值相较于二分类交叉熵损失 BCE 和 Dice 损失函数分别提高 1.97% 和 2.09%。联合损失函数在其余各项评价指标上均有提升,其中 MIoU 值最高提升 0.9%,表明在联合损失函数的作用下能有效提升植株叶片分割模型的整体分割性能。

3 结 论

针对植株叶片分割中目标叶片位置、边缘等细节信息的损失问题,在卷积神经网络中引入频域特征表示,设计下采样频域变换模块,加入改进后的通道注意力模块,构建空频域特征多尺度融合网络,采用联合损失函数引导网络的参数优化。在公开图像数据集 2017CVPPP 上的实验结果表明,本文植株叶片分割网络在交叉比、平均交叉比、像素准确率、精准率和 F-score 分数分别达到了 97.07%、98.04%、99.53%、99.68% 和 99.74%。与 FCN-8s、FCN-ResNet、DeepLabV3+、SegNet 和 U-Net 模型相比,在较小尺度和细节处理上改善了植株叶片的分割精度,为植物表型方向的应用研究奠定了基础。在后续的研究中将对其他频域特征表示做进一步的研究,在保证网络模型分割速度和精度的基础上提高效率。

参考文献

- [1] PIERUSCHKA R, SCHURR U. Plant phenotyping: Past, present, and future [J]. Plant Phenomics, 2019, 2019(3): 1-6.
- [2] FUENTES-PACHECO J, TORRES-OLIVARES J, ROMAN-RANGEL E, et al. Fig plant segmentation from aerial images using a deep convolutional encoder-decoder network[J]. Remote Sensing, 2019, 11(10): 1157.
- [3] SCHARR H, MINERVINI M, FRENCH A P, et al. Leaf segmentation in plant phenotyping: A collation

- study[J]. *Machine Vision and Applications*, 2015, 27(4): 585-606.
- [4] MINERVINI M, FISCHBACH A, SCHARR H, et al. Finely-grained annotated datasets for image-based plant phenotyping [J]. *Pattern Recognition Letters*, 2016, 81: 80-89.
- [5] AL-SHAKARJI N M, KASSIM Y M, PALANIAPPAN K. Unsupervised learning method for plant and leaf segmentation[C]. *2017 IEEE Applied Imagery Pattern Recognition Workshop(AIPR)*, 2017: 1-4.
- [6] KUMAR J P, DOMNIC S. Image based leaf segmentation and counting in rosette plants [J]. *Information Processing in Agriculture*, 2019, 6(2): 233-246.
- [7] HAMUDA E, GLAVIN M, JONES E. A survey of image processing techniques for plant extraction and segmentation in the field [J]. *Computers and Electronics in Agriculture*, 2016, 125: 184-199.
- [8] KUZNICHOV D, ZVIRIN A, HONEN Y, et al. Data augmentation for leaf segmentation and counting tasks in rosette plants[C]. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2019: 2580-2589.
- [9] AICH S, STAVNESS I. Leaf counting with deep convolutional and deconvolutional networks[C]. *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, 2017: 2080-2089.
- [10] 廖娟, 陈民慧, 张镔, 等. 基于区域语义和边缘信息融合的作物苗期植株分割模型[J]. *农业机械学报*, 2021, 52(12): 171-181.
- [11] 赵文慧, 杨霄, 孟丽洁. 充分利用多尺度特征改进 UNet 实现皮肤病变分割[J]. *电子测量技术*, 2022, 45(2): 110-116.
- [12] 钟思华, 王梦璐, 郭兴明, 等. 基于改进 VNet 的肺结节分割方法研究 [J]. *仪器仪表学报*, 2020, 41(9): 206-215.
- [13] ADAM D, JOHN P, SANJAY K, et al. Band-limited training and inference for convolutional neural networks [C]. *Proceedings of the 36th International Conference on Machine Learning*, 2019: 1745-1754.
- [14] PRATT H, WILLIAMS B, COENEN F, et al. FCNN: Fourier convolutional neural networks [J]. *Machine Learning and Knowledge Discovery in Databases*, 2017: 786-798.
- [15] AYAT S O, KHALIL-HANI M, RAHMAN A A-H A, et al. Spectral-based convolutional neural network without multiple spatial-frequency domain switchings [J]. *Neurocomputing*, 2019, 364: 152-167.
- [16] POUR M P, SEKER H. Transform domain representation-driven convolutional neural networks for skin lesion segmentation[J]. *Expert Systems with Applications*, 2020, 144: 113129.
- [17] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [J]. *Medical Image Computing and Computer-Assisted Intervention*, 2015: 234-241.
- [18] WU C Y, ZAHEER M, HU H, et al. Compressed video action recognition [C]. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018: 6026-6035.
- [19] XU K, QIN M, SUN F. et al. Learning in the frequency domain[C]. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 2020: 1737-1746.
- [20] ZHOU Z, SIDDIQUEE M M R, TAJBAKHS N, et al. UNet++: A nested U-net architecture for medical image segmentation [J]. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 2018: 3-11.
- [21] HE K, ZHANG X, REN S. et al. Identity mappings in deep residual networks [J]. *Computer Vision*, 2016: 630-645.
- [22] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [J]. *Computer Vision*, 2018: 3-19.
- [23] MILLETARI F, NAVAB N, AHMADI S A. V-Net: Fully convolutional neural networks for volumetric medical image segmentation [C]. *2016 Fourth International Conference on 3D Vision (3DV)*, 2016: 565-571.
- [24] LONG J, SHEHMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C]. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015: 3431-3440.
- [25] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [J]. *Computer Vision*, 2018: 833-851.
- [26] 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.
- [27] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks [C]. *2018 IEEE/CVF Conference*

on Computer Vision and Pattern Recognition(CVPR), 2018; 7132-7141.

- [28] YU T, LI X, CAI Y, et al. S^2 -MLPv2: Improved spatial-shift MLP architecture for vision[C]. 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2022; 3615-3624.

作者简介

罗薇, 硕士, 主要研究方向为计算机视觉、图像分割及分类相关应用。

E-mail: luowei@mail. ynu. edu. cn

陈峰, 硕士, 主要研究方向为烟叶生产。

E-mail: 250509935@qq. com

张晓伟, 硕士, 主要研究方向为烟叶质量。

E-mail: 52109672@qq. com

梁虹, 硕士, 教授, 主要研究方向为深度学习和信号处理。

E-mail: liangh@ynu. edu. cn

李军营(通信作者), 博士, 主要研究方向为烟草栽培。

E-mail: lji1250@163. com