

交互式双分支特征融合的草莓病害程度快速诊断方法

胡晓波^{1,2} 许桃胜² 黄伟² 王儒敬^{2,3}

(1. 中国科学技术大学研究生院, 合肥 230026; 2. 中国科学院合肥物质科学研究院, 合肥 230031;
3. 安徽省智慧农业工程实验室, 合肥 230031)

摘要: 针对现有草莓病害程度诊断方法存在识别精度低、参数量大、推理时间长等问题,提出了一种基于交互式双分支特征融合的草莓病害程度快速诊断方法。该方法首先以短程密集连接模块为基础,构建一种轻量化的交互式双分支特征融合网络(Interactive bilateral feature fusion network, IBFFNet),用于提取图像的语义特征和细节特征。然后,通过注意力简化的金字塔池化模块获取上下文分支中的多尺度语义特征,利用边缘增强模块丰富空间分支中的边缘细节特征。最后,融合多尺度语义特征和空间细节特征,实现病斑和叶片区域的精确分割。在草莓叶部病害程度数据集上的实验结果显示,IBFFNet2_Seg的平均交并比达到77.8%,在单张NVIDIA GTX1050显卡上处理速度可达40.6 f/s,满足实际应用中对于算法实时性和分割精度的要求。此外,在测试集上IBFFNet2_Seg预测病害程度与真实程度的决定系数 R^2 为0.98,说明该模型可以准确预测草莓病害严重程度。本研究可为草莓病害精准防治提供可靠的技术支撑。

关键词: 草莓病害; 语义分割; 病害程度快速诊断; 交互式双分支特征融合; 边缘增强

中图分类号: S431.9; TP391.4 文献标识码: A 文章编号: 1000-1298(2023)11-0225-11 OSID: 

Interactive Bilateral Feature Fusion Network for Real-time Strawberry Disease Diagnosis

HU Xiaobo^{1,2} XU Taosheng² HUANG Wei² WANG Rujing^{2,3}

(1. Graduate School, University of Science and Technology of China, Hefei 230026, China
2. Hefei Institutes of Physical Science, Chinese Academy of Science, Hefei 230031, China
3. Intelligent Agriculture Engineering Laboratory of Anhui Province, Hefei 230031, China)

Abstract: Accurately identifying the severity of strawberry leaf disease is essential for precise disease control. However, methods based on image classification had a rough division of disease severity and fuzzy classification boundary, while methods based on semantic segmentation had high computational costs and long inference time. To address these problems, a real-time strawberry disease diagnosis method was proposed based on interactive bilateral feature fusion network (IBFFNet). The IBFFNet was a lightweight model containing a context path and a spatial path to extract semantic and detail features from the input image, respectively. Furthermore, an attention spatial pyramid pooling module was constructed to extract multi-scale semantic features from the context path, and an edge enhancement module was designed to enrich edge detail information in the spatial path. Finally, the multi-scale semantic feature and detail information were aggregated for precise leaf and lesion area segmentation. The percentage of lesions in the leaf area was the estimated severity. The method achieved a promising trade-off between accuracy and speed on the strawberry leaf disease diagnosis dataset. On the strawberry leaf disease diagnosis dataset, the mIoU of IBFFNet2_Seg was 77.8% with 40.6 f/s on a single NVIDIA GTX1050. In the test set, an R^2 value (coefficient of determination) of 0.98 was achieved, which denoted that the IBFFNet2_Seg could accurately predict the severity of the three diseases. This study paved the way for the precise control of strawberry disease.

Key words: strawberry disease; semantic segmentation; real-time disease severity diagnosis; interactive bilateral feature fusion; edge enhancement

收稿日期: 2023-09-05 修回日期: 2023-09-28

基金项目: 国家重点研发计划项目(2019YFE0125700、2021YFD2000201)

作者简介: 胡晓波(1988—),男,博士生,主要从事深度学习与农作物病虫害识别等研究,E-mail: huxiaobo@mail.ustc.edu.cn

通信作者: 王儒敬(1964—),男,研究员,博士,主要从事智慧农业、农业大数据和数据挖掘等研究,E-mail: rjwang@iim.ac.cn

0 引言

中国是全球最大的草莓生产国^[1]。据国家统计局数据显示,2021 年我国草莓种植面积约为 $1.4 \times 10^5 \text{ hm}^2$,年产量达 $3.7 \times 10^6 \text{ t}$ 。草莓种植业已成为我国多个市县的支柱性产业。但在草莓的整个栽培过程中易受各类病原体侵染,这对草莓产量和品质产生了重要影响^[2]。病害已成为影响我国草莓产业持续健康发展的主要因素。在发现草莓病害后,种植户通常会在病株上喷洒杀菌剂,以防止病害的进一步发展和蔓延。然而,剂量不足可能无法有效控制病害,过量则可能导致农残超标和环境污染。因此,快速准确地识别草莓病害严重程度,实现病害精准防治,对降低种植风险、提高草莓品质具有重要意义。

当前,我国草莓病害程度诊断主要依赖于人工田间调查,这种方式对植保人员的专业知识要求较高,且存在成本高、实时性差、主观性强等缺点^[3-5]。近年来,随着深度学习技术在语音识别、无人驾驶和医疗诊断等领域取得巨大成功,研究人员开始探索利用深度学习技术实现病害程度的自动诊断^[6-13]。基于深度学习的病害程度诊断方法根据实现手段可分为基于图像分类方法和基于语义分割方法。

基于图像分类的作物病害程度诊断方法将病害程度诊断问题转化为图像分类问题,在病害类别的基础上根据病害程度对同一类别病害作进一步划分,并利用分类算法实现病害程度诊断^[5-7,9,13]。杜甜甜等^[5]将作物病害程度划分为健康、一般和严重 3 个等级,利用通道注意力机制和多尺度特征融合策略,提升模型对同类病害不同程度的分类精度。何东健等^[6]将葡萄霜霉病按严重程度划分为健康、轻度、中度和重度 4 个等级,提出一种改进的残差网络对不同的病害程度进行诊断,该方法平均识别精度为 99.92%。于明等^[7]通过构建颜色掩膜网络和通道自适应注意力机制,增强了卷积神经网络提取病斑颜色、位置、上下文等特征的能力,增加病斑特征在病害程度诊断时的权重,其在 AI Challenger 2018 数据集上的 Top 1 准确率达 87.97%。这些研究表明利用图像分类技术识别作物病害程度在技术上是可行的。然而,这种方法对病害程度的划分略显粗糙(最多分为健康、轻度、中度和重度 4 个级别),且不同程度间的分类边界模糊,缺乏准确、有效的度量标准,很难为病害防治提供精准指导,实际应用价值有限。

基于语义分割的作物病害程度诊断方法利用图像分割算法将叶片和病斑分别从图像中提取出来,

通过计算病斑与叶片面积比得到准确的病害程度值^[10-12]。WSPANIALY 等^[10]从 PlantVillage 数据集中选取 9 种番茄病害的 1 800 幅图像做像素级标注,建立所需的番茄病害程度数据集,使用 U-Net^[14]对病斑和叶片进行分割,通过计算病斑与叶片面积比得到病害程度,其病害程度识别平均误差为 11.86%。GONÇALVES 等^[11]选取咖啡叶矿虫损伤、大豆锈病、小麦棕斑病为研究对象,测试 5 个经典分割网络(UNet^[14]、SegNet^[15]、PSPNet^[16]、FPN^[17]和 DeepLab V3 +^[18])的病害程度诊断性能,结果表明 UNet^[14]、FPN^[17]和 DeepLab V3 +^[18]的病害程度估计准确率接近于人工 APS 评估分析结果。与人工方法相比,语义分割方法的速度更快、效率更高,更适合应用于生产实践。因此,利用语义分割方法能够更加精确地预测作物的病害程度,可以为作物病害防治提供更加可靠的数据支撑。然而,上述研究中所提出的病害程度诊断方法存在模型参数量大、病害推理时间长等缺点,并不适合移动端部署,不能满足病害程度现场实时诊断的需求。综上所述,研究轻量化的实时病害程度诊断方法是非常有必要的。

针对上述问题,本文在短程密集连接(Short-term dense concatenate, STDC)模块的基础上,构建一种轻量化的交互式双分支特征融合网络,分别提取图像的语义特征和细节特征。此外,设计注意力简化的金字塔池化模块获取上下文分支中的多尺度语义特征,利用边缘增强模块丰富空间分支中的边缘细节特征。最后,融合多尺度语义特征和边缘细节特征,实现对叶片和病斑区域的分割。病斑与叶片的面积比是最终的病害程度诊断结果。

1 材料与方法

1.1 图像采集与数据集建立

考虑到病害发生率和病斑区域的可标记性,选择草莓白粉病、褐斑病和蛇眼病为研究对象,开展草莓病害程度快速诊断方法研究。图像采集地点为安徽省长丰县智慧农业谷现代农业草莓园(31°97'N, 117°28'E)。采集时间为 2022 年 9 月至 2023 年 3 月,采集设备为红米 note 11 和荣耀 v 40 智能手机,图像分辨率最大为 3 042 像素 × 4 032 像素。为保证病害程度诊断的准确性,采集时将手机摄像头置于叶片正上方 15 ~ 20 cm 处,使整个叶片包含在图像中。对采集后的图像进行筛选,去除重复、叶片遮挡、卷曲、残缺的图像,得到 3 类草莓病害图像各 200 幅。通过 Lableme 软件对图像中的病斑、叶片和背景区域进行标注,构建用于本文研究的草莓叶片

病害程度诊断数据集 (Strawberry leaf disease diagnosis dataset, SLDDD)。数据集典型示例如图 1 所示。

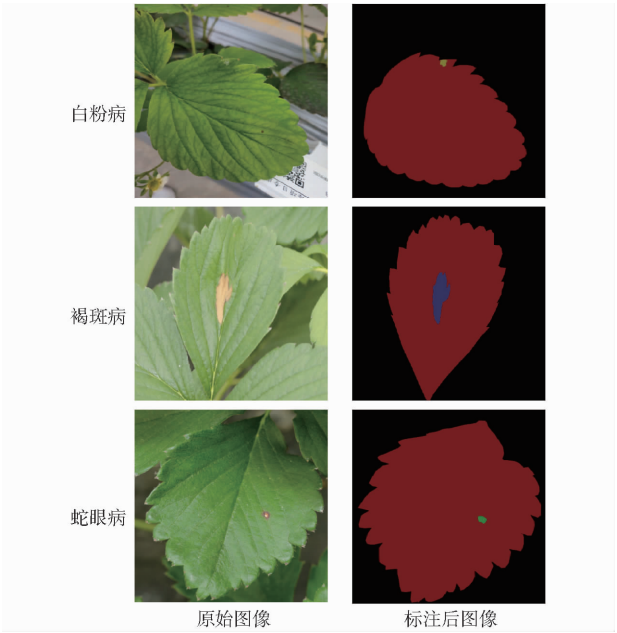


图 1 SLDDD 数据集示例

Fig. 1 Examples of SLDDD

1.2 交互式双分支特征融合的草莓病害程度快速诊断方法

基于卷积神经网络的语义分割方法根据推理速度可分为两类。一类是采用编码-解码结构的高精度语义分割方法,其中编码网络负责提取高级语义特征,解码网络负责将高级语义特征与不同层的细节特征依次融合,以不断提升网络的分割精度。尽管,这类模型的预测精度高,但计算量大、推理速度慢,代表性模型有 U-Net^[14] 和 DeepLab V3 +^[18] 等。另一类是基于双分支结构的实时语义分割模型,其中一个分支负责提取图像中的上下文语义特征,另一个分支负责提取空间细节特征。最后,通过两分支的特征融合实现图像分割,其经典模型有 BiSeNetV1^[19]、BiSeNetV2^[20]、STDC_Seg^[21]、PP-

LiteSeg^[22] 和 DDRNet_Seg^[23] 等。本文在双分支语义分割方法的基础上,通过两分支间的特征交互融合,不断丰富上下文分支中的细节特征、增强空间分支中的语义特征。此外,构建通道注意力简化的特征金字塔模块获取上下文分支中的多尺度语义特征,设计边缘增强模块丰富空间分支中的细节特征。

1.2.1 交互式双分支特征融合网络

为减小模型参数数量和计算时间,在文献[21]所提出的 STDC 模块基础上,构建一种轻量的交互式双分支特征融合网络 IBFFNet,提取图像中的语义特征和细节特征。网络架构如图 2 所示。

除输入、输出外,IBFFNet 由 6 个阶段组成。其中,阶段 1~5 以步距 2 缩减输入图像空间尺寸,阶段 6 通过一个全局平均池化和一个全连接层得到网络的预测概率值。此外,在 1/4 图像分辨率(阶段 2)处开始构建空间分支,在 1/8 和 1/16 图像分辨率(阶段 3 和阶段 4)处通过双向特征融合模块(Bidirectional feature fusion module, BFFM)实现两分支中的上下文特征和细节特征的交互融合。图中的 ConvBlock 表示由一个卷积层、一个批归一化层和一个 ReLU 激活函数组成的卷积模块。STDC 模块和 BFFM 是网络的关键组件。

表 1 列出了 IBFFNet1 和 IBFFNet2 的网络结构,其中 K、C、S、R、O 和 d 分别表示卷积核、特征通道数、步距、模块重复次数、输出特征图尺寸和输出神经元个数。IBFFNet 的上下文分支与文献[21]所提出的 STDC 模型的差别在于:从阶段 2 开始加入 STDC 模块,以增大该阶段网络的感受野,增强网络感知语义特征的能力。去除 STDC 模型最后一个卷积层和全连接层,进一步缩减网络的参数数量和计算量。IBFFNet1 和 IBFFNet2 是本文所构建 2 种交互式双分支特征融合网络,二者网络结构相同,只是在各阶段中 STDC 模块的重复数量以及空间分支特征通道数上存在差别。

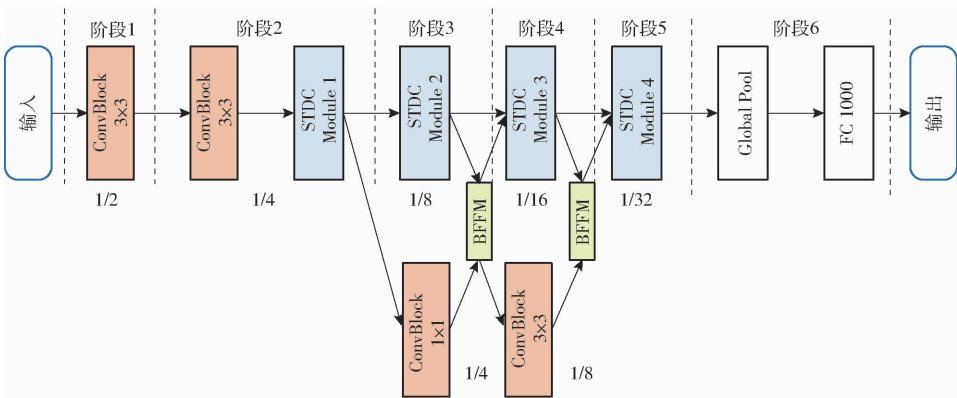


图 2 IBFFNet 结构图

Fig. 2 Framework of IBFFNet

表 1 IBFFNet 网络的详细结构

Tab.1 Details of IBFFNet				
IBFFNet1			IBFFNet2	
阶段 1	ConvBlock $K=3\times3$		ConvBlock $K=3\times3$	
	$C=32, S=2, R=1$		$C=32, S=2, R=1$	
	$O=112\times112$		$O=112\times112$	
阶段 2	ConvBlock $K=3\times3$		ConvBlock $K=3\times3$	
	$C=64, S=2, R=1$		$C=64, S=2, R=1$	
	$O=56\times56$		$O=56\times56$	
	STDC 模块		STDC 模块	
	$C=128, S=1, R=1$		$C=128, S=1, R=1$	
$O=56\times56$			$O=56\times56$	
阶段 3	STDC 模块		STDC 模块	
	ConvBlock		$C=256,$	
	$K=1\times1$		$S=1, R=1$	
	$C=256,$		$K=1\times1$	
	$S=1, R=1$		$O=28\times28$	
	$O=28\times28$		STDC 模块	
			$C=128,$	
		$S=1, R=1$		
		$O=56\times56$		
		$S=1, R=3$		
		$O=28\times28$		
BFFM			BFFM	
阶段 4	STDC 模块		STDC 模块	
	$C=512,$		$C=512,$	
	$S=2, R=1$		$S=2, R=1$	
	ConvBlock		ConvBlock	
	$K=3\times3$		$K=3\times3$	
	$O=14\times14$		$O=14\times14$	
	STDC 模块		STDC 模块	
	$C=512,$		$C=512,$	
$S=1, R=1$		$S=1, R=4$		
$O=14\times14$		$O=28\times28$		
		$O=14\times14$		
BFFM			BFFM	
阶段 5	STDC 模块		STDC 模块	
	$C=1\ 024, S=2, R=1$		$C=1\ 024, S=2, R=1$	
	$O=7\times7$		$O=7\times7$	
	STDC 模块		STDC 模块	
	$C=1\ 024, S=1, R=1$		$C=1\ 024, S=1, R=2$	
	$O=7\times7$		$O=7\times7$	
阶段 6	全局平均池化		全局平均池化	
	$C=1\ 024, O=1\times1$		$C=1\ 024, O=1\times1$	
	全连接, Softmax		全连接, Softmax	
	$1000-d$		$1000-d$	

STDC 模块网络结构如图 3 所示,一个 STDC 模块由多个 ConvBlock 组成。其中第 1 个 ConvBlock 的卷积核为 1,用于特征通道变换。后面的 ConvBlock 的卷积核均为 3,各 ConvBlock 的输出被

连接(Concatenate)在一起作为整个模块的最终输出。

假设一个 STDC 模块有 n 个 ConvBlock,输入、输出特征通道数分别为 M 和 N ,则该模块的第 i 个 ConvBlock 输出通道数为

$$C_i=\begin{cases}\frac{N}{2^i} & (i=1,2,\cdots,n-1) \\ \frac{N}{2^{i-1}} & (i=n)\end{cases}\tag{1}$$

显然,随着 STDC 模块中 ConvBlock 层数的增加,ConvBlock 的输出通道数会以 2 的指数幂递减。整个模块参数量计算公式为

$$\begin{aligned}S_{params} &= M\frac{N}{2^1}\times1\times1+\sum_{i=2}^{n-1}\left(\frac{N}{2^{i-1}}\frac{N}{2^i}\times3\times3\right)+ \\ &\frac{N}{2^{n-1}}\frac{N}{2^{n-1}}\times3\times3= \\ &\frac{MN}{2}+\frac{9N^2}{2^3}\sum_{i=0}^{n-3}\frac{1}{2^{2i}}+\frac{9N^2}{2^{2n-2}}= \\ &\frac{N(M+3N)}{2}+\frac{3N^2}{2^{2(n-1)}}\end{aligned}\tag{2}$$

从式(2)可以看出,模块参数量主要取决于输入通道数 M 和输出通道数 N ,层数 n 对参数量的影响很小。因此,模块参数量和计算量不会随着模块深度的增加而明显增加。然而,增加 ConvBlock 层数有利于扩大模块感受野,提升上下文分支获取语义信息的能力。本文 STDC 模块中 n 默认设置为 4^[21]。

双向特征融合模块主要用于融合上下文分支的语义特征和空间分支的细节特征,通过两分支间的特征互补,增强模型整体分割效果。该模块结构如图 4 所示。由于上下文语义特征 $F_{Semantic}$ 和空间细节特征 F_{Detail} 的通道数、分辨率不同,在融合前需进行通道变换和尺寸缩放。对于 $F_{Semantic}$ 融入 F_{Detail} ,BFFM 先通过注意力精炼模块(Attention refinement module,ARM)对 $F_{Semantic}$ 的通道数进行压缩,再通过双线性插值算法将压缩后的特征上采样至 F_{Detail} 的尺寸。对于 F_{Detail} 融入 $F_{Semantic}$,BFFM 直接通过一个核为 3×3 、步长为 2 的卷积将 F_{Detail} 的通道数和尺寸变换成与 $F_{Semantic}$ 相同。在实现 $F_{Semantic}$ 和 F_{Detail} 的通

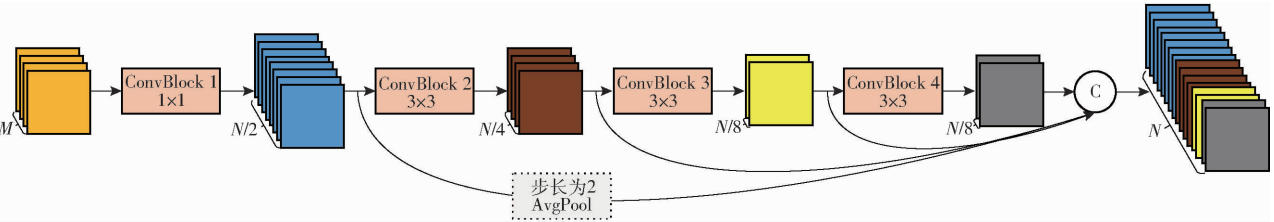


图 3 STDC 模块结构

Fig.3 Architecture diagram of STDC module

道数、分辨率的对齐后,BFFM 直接通过特征对应元素相加实现两分支间的特征融合。

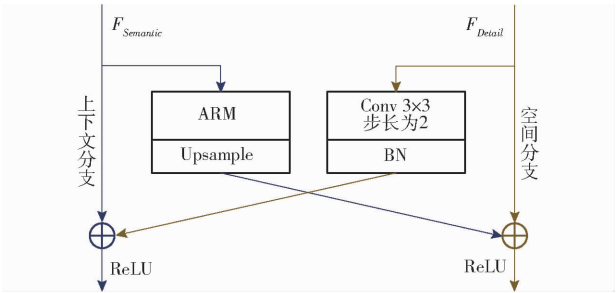


图 4 BFFM 结构图
Fig.4 Structure diagram of BFFM

文献[19,24]指出特征通道数压缩会造成信息丢失,为此在进行通道压缩前有必要评估每个通道的重要性,并增强重要性高的通道,抑制冗余通道。采用注意力精炼(ARM)模块,在通道变换前通过通道注意力机制实现重要性通道增强和冗余通道抑制,再由一个 1×1 卷积实现通道变换。模块结构如图 5 所示。

1.2.2 病害程度快速诊断网络

使用在 ImageNet 数据集上预训练的 IBFFNet,作为病害快速诊断模型的骨干网络,编码图像中的语义特征和细节特征。如图 6 所示,在空间分支中添加边缘增强模块 (Edge enhancement module, EEM) 丰富空间分支中的边缘细节特征,在上下文

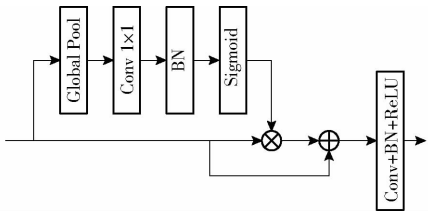


图 5 ARM 结构图
Fig.5 Structure diagram of ARM

分支中加入注意力简化的金字塔池化模块 (Attention simplified pyramid pooling module, ASPPM) 增强模型获取多尺度语义特征的能力。最后,通过对应元素相加实现多尺度语义特征和细节特征的融合。图 6 中,SegHead 层用于从融合特征中得到预测结果。本文的 SegHead 层由一个 3×3 的 ConvBlock 和一个 1×1 的 ConvBlock 组成,其输出特征通道数是预测类别数 N_c 。

文献[25]指出增强特征中的边缘信息是提高物体边缘分割精度的一种有效手段。对于田间复杂环境下的叶片分割,由于背景中时常存在其他叶片干扰,使叶片与背景的对比度非常小,因此在空间分支中增强边缘特征将有助于提高模型对叶片和病斑边缘的分割精度。

通过构建边缘增强模块 EEM,实现对细节分支中边缘特征的增强,其结构如图 7 所示。EEM 首先利用一个 1×1 卷积、批归一化层和 Sigmoid 函数预

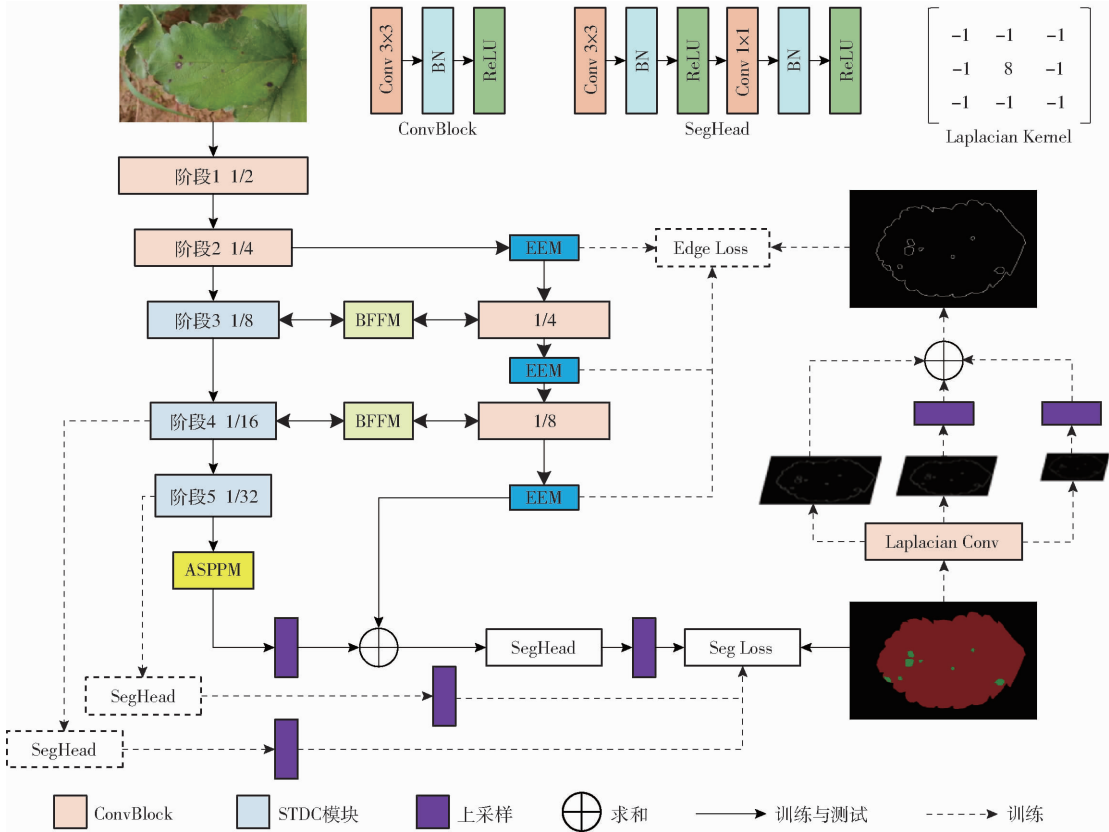


图 6 基于交互式特征融合的病害程度诊断模型架构图

Fig.6 Framework of real-time strawberry disease severity diagnosis model based on interactive bilateral feature fusion

测空间细节特征 F_{Detail} 中叶片和病灶的边缘,再通过边缘与 F_{Detail} 相乘后相加实现边缘特征增强,其流程可概括为

$$F_{Enhanced} = F_{Detail}(1 + p_e) \tag{3}$$

其中 $p_e = \text{Sigmoid}(\text{BN}(\text{Conv}(F_{Detail})))$

式中 p_e ——预测边缘

$F_{Enhanced}$ ——边缘增强后特征

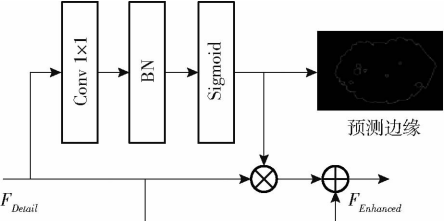


图 7 EEM 结构图

Fig. 7 Structure diagram of EEM

此外,定义边缘损失用于模块预测边缘的优化,以确保 EEM 具备边缘提取功能。由于物体边缘像素在整幅图像中的比例很小,因此边缘预测是一个类别极度不平衡的问题,使用加权交叉熵损失并不能很好解决该问题^[25]。本文采用二值交叉熵和 Dice 损失来优化 EEM 对物体边缘的学习^[21,25],即

$$L_{edge}(p_e, g_e) = L_{bce}(p_e, g_e) + L_{dice}(p_e, g_e) \tag{4}$$

其中

$$L_{bce}(p_e, g_e) = \frac{1}{HW} \sum_{i=1}^{HW} [p_e^i \lg g_e^i + (1 - p_e^i) \lg (1 - g_e^i)] \tag{5}$$

$$L_{dice}(p_e, g_e) = 1 - \frac{2 \sum_{i=1}^{HW} p_e^i g_e^i + \varepsilon}{\sum_{i=1}^{HW} (p_e^i)^2 + \sum_{j=1}^{HW} (g_e^j)^2 + \varepsilon} \tag{6}$$

式中 L_{edge} ——边缘损失值

g_e ——真实边缘

L_{bce} ——二值交叉熵损失值

L_{dice} ——二值 Dice 损失值

ε ——避免除零的拉普拉斯平滑项

文献[16]所提出的特征金字塔池化模块(Pyramid pooling module, PPM)可以有效获取多尺度全局上下文信息。对于草莓病害,部分病斑的尺度小、形态多变,采用金字塔池化将有助于提升模型对 these 小病斑区域的分割精度。但是,上下文分支最后一层的通道维度一般比较高,通过特征连接操作融合不同尺度特征,会使特征维度成倍增加,导致模型参数量和计算量激增。文献[22]通过缩减 PPM 中间层层数和通道数,并将不同尺度特征间的融合操作由连接改为对应元素相加,提出一种简化

的特征金字塔池化模块(Simplified pyramid pooling module, SPPM)。然而,SPPM 没有融合原始特征图,也未考虑通道间的关联性,致使其对模型分割性能提升效果有限。

在文献[22]的基础上构建注意力简化的特征金字塔模块 ASPPM,其结构如图 8 所示。ASPPM 首先使用 3 个全局平均池化操作,将原特征图的空间尺寸分别变换为 1×1 、 2×2 和 4×4 ;再通过 1.2.1 节所提出的 ARM 模块缩减这些特征的通道数,并将缩减后的特征图上采样至原特征图尺寸;最后,通过求和操作实现不同尺度特征的融合。

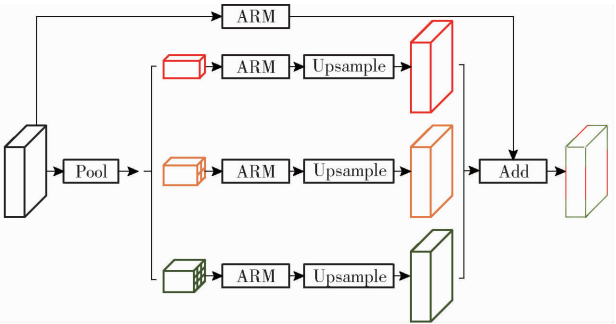


图 8 ASPPM 结构图

Fig. 8 Structure diagram of ASPPM

直接采用文献[21]所提出的边缘生成方法,即通过一个以拉普拉斯模板为权重的卷积层,以 3 种不同步距(1,2,4)从图像真实语义标签图中提取不同尺度的边缘图。再将这些提取的边缘图通过双线性插值,上采样至真实标签图尺寸。随后,通过一个可学习的 1×1 卷积实现多尺度边缘图的融合。最后,以阈值 0.1 从融合的边缘图中生成真实边缘。

2 结果与分析

2.1 实验环境与评价指标

2.1.1 实验数据集

在 ImageNet 数据集^[26](ILSVRC 2012)上训练并测试所提出的交互式特征融合网络(IBFFNet1 和 IBFFNet2)的分类性能。该数据集有 1 000 个类别,其训练集、验证集分别包含 1.2×10^6 幅和 5×10^4 幅图像。病害程度快速诊断模型在 1.1 节所构建的草莓叶片病害程度数据集(SLDDD)上训练和验证。由于 SLDDD 中每类病害样本量较少(每类仅有 200 个样本),实验采用 5 折交叉验证进行数据集划分与模型验证,即先从 SLDDD 每类病害中随机选择 50 个样本作为测试集,将剩余样本分为 5 等份(每份 30 个样本)。在测试集不变的情况下进行 5 次实验,每次从中选择 4 份作为训练集,1 份作为验证集(保证每次实验所用的验证集互不相同),验证集上表现最佳的模型在测试集中的预测结果作为该次实

验的结果。最后,取 5 次实验结果的平均值作为模型最终预测值。每次实验通过随机镜像、随机旋转和随机高斯模糊对训练集进行在线数据增强,以提高模型鲁棒性,防止模型过拟合。

2.1.2 实验设置与环境

采用随机梯度下降算法 (Stochastic gradient descent,SGD) 进行模型参数优化,其初始学习率、动量和权重衰减分别设置为 0.05、0.9 和 1×10^{-4} 。模型在 ILSVRC 2012^[26] 上训练 300 个迭代周期 (Epoch),前 5 个 Epoch 利用学习率预热策略将学习率从 0.001 线性提升至初始学习率,此后使用余弦策略更新 SGD 优化器的学习率。每批次训练 64 幅图像,每幅图像尺寸被缩放至 224 像素 $\times$ 224 像素。

采用 SGD 优化器优化模型参数,其初始学习率、动量和权重衰减分别设置为 0.01、0.9 和 5×10^{-4} 。模型在 SLDDD 数据集上迭代 5 000 次,前 500 次利用学习率预热策略将学习率从 0.000 1 线性提升至 0.01,此后使用“poly”策略更新 SGD 优化器的学习率,即每次迭代的学习率为 $\left(1 - \frac{i_{ter}}{5\,000}\right)^{p_{ower}}$ 。其中, i_{ter} 为当前迭代次数, p_{ower} 设置为 0.9。每批次训练 32 幅图像,每幅图像尺寸被缩放至 512 像素 $\times$ 512 像素。

本节所提出的算法均在 PyTorch 1.9.0 框架下实现,在一台硬件环境为 Xeon W - 2200 处理器、NVIDIA GeForce RTX 3090 GPU,软件环境为 CUDA 11.3、CUDNN 7.6.4、Ubuntu 18.04 操作系统的工作站上执行所有模型的训练和测试。此外,为方便与其他算法对比,使用一台搭载 CUDA 11.0、CUDNN 7.6.4、NVIDIA GTX1050 GPU 的计算机测试各模型的推理速度。

2.1.3 评价指标

采用 Top 1 精度作为分类模型的评价指标。Top 1 精度是模型正确预测的样本数占总样本数的比例。采用输入图像真实像素类别与模型预测像素类别的平均交并比 (Mean intersection over union, mIoU) 作为衡量模型分割精度的指标。此外,使用模型每秒处理图像的帧数作为模型推理速度 (效率) 的评价指标。

2.2 模型性能对比

2.2.1 分类性能对比

表 2 为 IBFFNet 与其他先进的轻量模型的识别精度、参数量、计算量和推理速度 4 个指标。所有模型均使用分辨率为 224 像素 $\times$ 224 像素的图像作为输入。尽管 IBFFNet 在 ImageNet 上的分类精度略低

于 STDC 模型^[21],但 IBFFNet 的参数量和计算量远低于 STDC 模型^[21],IBFFNet 的推理速度也快于 STDC。相比于其他模型,IBFFNet 在分类精度和推理速度上取得了更好的平衡。IBFFNet1 获得了最高的推理速度 189.4 f/s;在识别精度大体相同的情况下,IBFFNet2 较 DDRNet23^[23] 参数量、计算量分别减少 1.7×10^7 和 2.0×10^9 ,推理速度提高 32.8 f/s。

表 2 IBFFNet 与先进轻量化模型在 ImageNet 验证集上的分类性能对比

Tab.2 Comparison results of IBFFNet and some state-of-the-art lightweight models on validation set of ImageNet

模型	Top 1 精度/%	参数量	计算量	推理速度/ (f·s ⁻¹)
ResNet - 18	69.0	1.1×10^7	1.8×10^9	172.2
ResNet - 50	75.3	2.4×10^7	3.8×10^9	62.9
MobileNetV2	72.0	3.4×10^6	3.0×10^8	118.6
MobileNetV3	75.2	5.4×10^6	2.2×10^8	99.8
DDRNet23slim	70.2	7.6×10^6	1.0×10^9	111.8
DDRNet23	75.9	2.8×10^7	4.0×10^9	67.3
STDC1	73.9	1.4×10^7	3.2×10^9	156.1
STDC2	76.4	1.9×10^7	5.1×10^9	86.2
IBFFNet1	72.6	6.8×10^6	9.9×10^8	189.4
IBFFNet2	76.1	1.1×10^7	2.0×10^9	100.1

2.2.2 病害分割性能对比

表 3 比较了多个先进的实时语义分割模型在 SLDDD 测试集上的语义分割性能。所有模型均在 ImageNet 数据集上进行预训练,并使用分辨率为 512 像素 $\times$ 512 像素的图像作为输入。与其他实时语义分割网络相比,本节所提出的模型在分割精度 (mIoU)、参数量和推理速度上得到了更好的权衡。具体地说,IBFFNet1_Seg 获得了最高的推理速度 61.9 f/s,且分割精度达到 74.4% (较 BiSeNetV1^[19] 高 1.9 个百分点);IBFFNet2_Seg 在 SLDDD 测试集

表 3 IBFFNet_Seg 与先进实时语义分割模型在 SLDDD 测试集上的分割性能比较结果

Tab.3 Comparison results of IBFFNet_Seg and some state-of-the-art real-time segmentation models on test set of SLDDD

模型	骨干网络	mIoU/%	参数量	推理速度/ (f·s ⁻¹)
BiSeNetV1	ResNet - 18	72.5	1.3×10^7	51.6
BiSeNetV2		71.4	5.2×10^6	37.1
STDC1_Seg	STDC1	75.3	1.4×10^7	49.7
STDC2_Seg	STDC2	76.9	2.2×10^7	40.0
DDRNet_Seg	DDRNet23	76.4	2.0×10^7	38.6
IBFFNet1_Seg	IBFFNet1	74.4	1.1×10^7	61.9
IBFFNet2_Seg	IBFFNet2	77.8	1.6×10^7	40.6

上获得了最高的分割精度 77.8% ,且推理速度快于 STDC2_Seg^[21] 和 DDRNet_Seg^[23]。

2.3 消融实验

消融实验主要用于验证所构建的病害程度快速诊断模型中各模块的有效性。

2.3.1 BFFM

BFFM 的设计目的是实现上下文分支语义特征和空间分支细节特征的双向融合。为验证 BFFM 的有效性,表 4 比较了无特征交互模块、使用文献[23]所提出的双向融合模块 (Bilateral fusion module, BFM) 以及本文所提出的 BFFM 情况下模型的分割性能。从表 4 中可以看出,语义特征和细节特征的交互有助于模型分割精度的提升,如在 IBFFNet1 中 BFM、BFFM 较无交互的模型分割性能分别提升 1.4、2.3 个百分点。此外,相较于 BFM,使用本节所提出的 BFFM,IBFFNet1 和 IBFFNet2 的分割精度分别提高 0.9、1.2 个百分点,说明本文所提出的 BFFM 具有更好的特征交互融合效果。

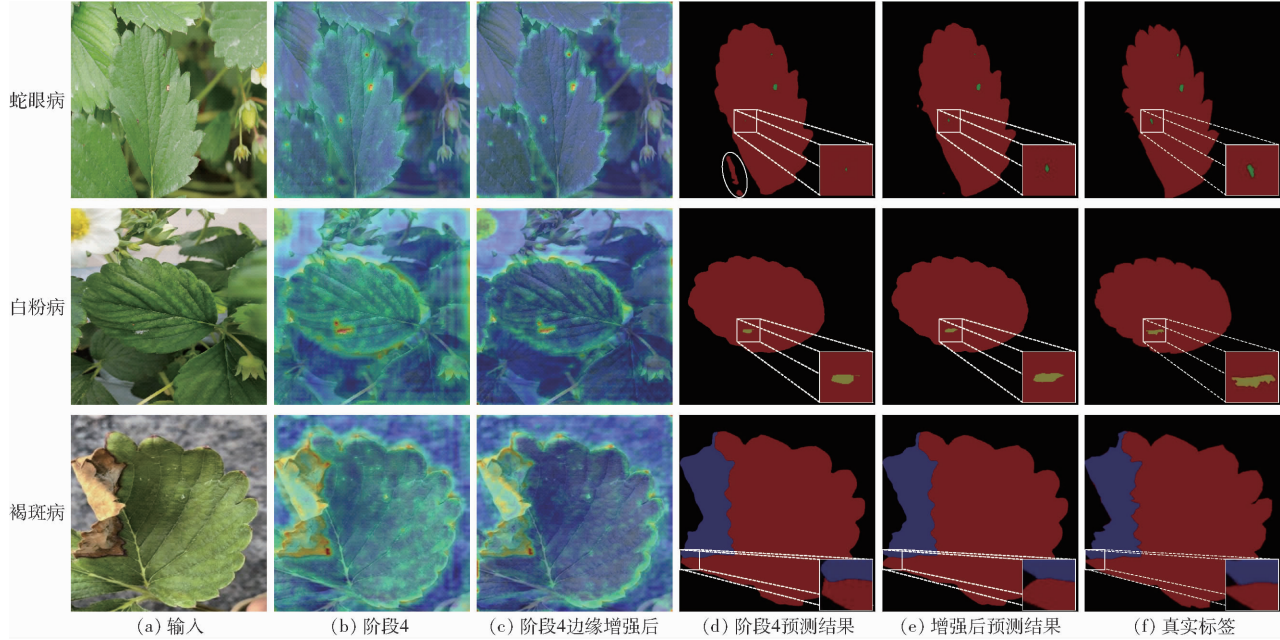


图 9 EEM 边缘增强效果可视化比较

Fig.9 Visual comparisons of edge enhancement effects in EEM

EMM 后细节特征集中于主叶片边缘及其病灶区域,其他部位细节特征被有效抑制。这表明 EMM 增强了叶片、病灶边缘与背景的对比度,提高了模型对叶片和病灶边缘的分割精度。

表 5 展示了不同阶段加入 EEM 模块,模型预测精度的变化情况。可以看出,EEM 以牺牲很小的推理速度为代价,提升了模型分割精度。此外,越深层的特征中加入 EEM,越有助于模型预测精度的提升。因为,深层特征的语义信息强,所包含的边缘语义信息更加明确,边缘增强的效果也就越明显。

表 4 不同特征融合方法比较

Tab.4 Performance comparison of different features fusion methods

模型	特征交互方法		mIoU/ %	推理速度/ (f·s ⁻¹)
	BFM	BFFM		
IBFFNet1_Seg	√		72.1	77.4
		√	73.5	64.0
			74.4	61.9
IBFFNet2_Seg	√		75.0	50.9
		√	76.6	41.8
			77.8	40.6

注:“√”表示模型使用该模块。下同。

2.3.2 EEM

设计 EEM 的目的是增强空间分支中叶片和病斑的边缘特征。为验证 EEM 的有效性,将阶段 4 对应的细节分支特征图在边缘增强前后的热力图以及模型加入 EEM 前后的预测结果进行可视化,结果如图 9 所示。对比图 9b 和图 9c 可以发现,在加入 EEM 前细节特征集中于整幅图像中的物体,加入

2.3.3 ASPPM

为验证 ASPPM 的有效性,表 6 比较了在 IBFFNet2 中使用不同的空间金字塔模块获取多尺度上下文信息对模型分割性能的影响。相比于没有获取上下文特征的模型,加入 SPPM^[22]、PPM^[16] 和 ASPPM,模型分割精度分别上升 0.6、2.3、2.1 个百分点,说明获取多尺度上下文信息有助于模型分割精度的提升。尽管,使用 PPM^[16] 获取上下文信息的模型分割精度最高,但 PPM^[16] 的计算代价最大,造成模型推理速度下降 7.9 f/s,而 SPPM^[22] 和 ASPPM

表 5 IBFFNet2_Seg 不同阶段加入 EMM 分割性能比较
Tab.5 Performance comparison of IBFFNet2_Seg with EEM added at different stages

阶段 2	阶段 3	阶段 4	mIoU/%	推理速度/ (f·s ⁻¹)
			75.5	43.8
√			76.2	41.7
	√		76.5	41.7
		√	76.9	41.7
	√	√	77.3	41.1
√	√	√	77.8	40.6

表 6 不同多尺度上下文特征获取方法下 IBFFNet_Seg2 的分割性能比较

Tab.6 Performance comparison of IBFFNet_Seg2 under different multi-scale contextual feature aggregation methods

多尺度上下文获取方法			mIoU/%	推理速度/ (f·s ⁻¹)
PPM	SPPM	ASPPM		
			75.7	44.3
√			76.3	40.9
	√		78.0	36.4
		√	77.8	40.6

对推理速度的影响较小,下降幅度仅为 3.4 f/s 和

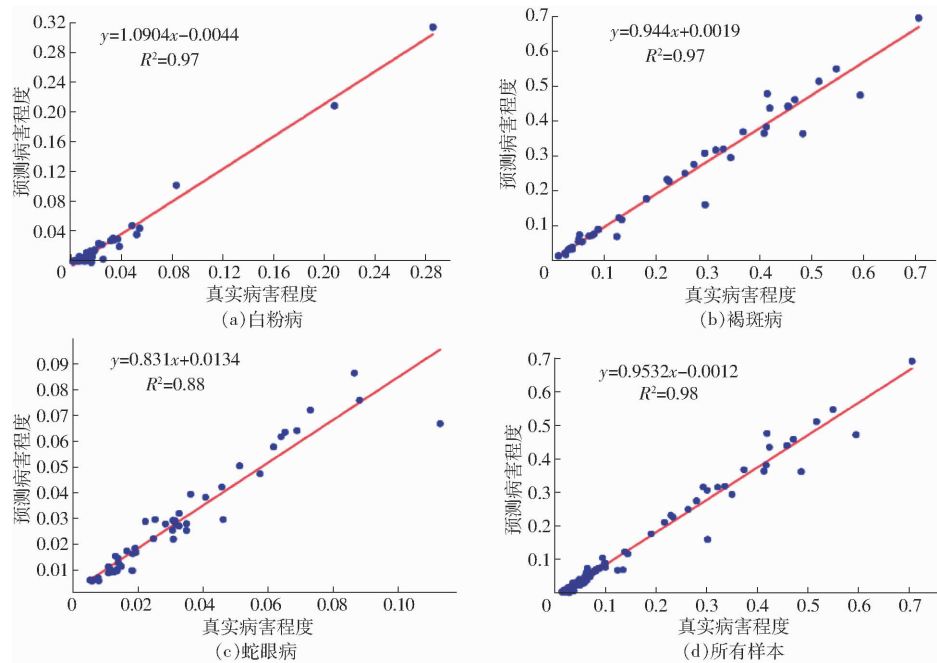


图 10 IBFFNet2_Seg 预测的病害程度与真实病害程度间的拟合关系曲线
Fig. 10 Liner relationships between true severity and severity predicted by IBFFNet2_Seg

3 结论

(1) 针对基于语义分割的病害程度诊断模型参数量大、推理速度慢的问题,提出了一种轻量化的交互式双分支特征融合网络 IBFFNet。由于采用短程密集连接方式,IBFFNet 可以在不显著增加模型参数数量的情况下,增加模型深度、扩大模型感

3.7 f/s。因此,相比于 SPPM^[22] 和 PPM^[16],本节所提出的 ASPPM 是一种简单、高效的获取多尺度上下文特征的方法。

2.4 病害程度估计

使用 SLDDD 的测试集验证 IBFFNet2_Seg 病害程度估计的有效性。利用每幅图像中标记的真实病斑区域与叶片区域的面积比作为真实病害程度值,IBFFNet2_Seg 预测的病斑区域与叶片区域的面积比作为病害程度估计值。白粉病、褐斑病、蛇眼病以及所有病害在一起,模型预测的病害程度与真实病害程度的拟合关系如图 10 所示。由于所采集的样本多数来自草莓种植园而不是试验田,中晚期病害特征明显,管理人员在发现后通常会将其摘除,导致所采集的样本中低程度病害样本较多(特别是白粉病和蛇眼病,病斑占叶片的比例多在 10% 以下)。此外,蛇眼病早期症状表现为暗紫红色小斑点(图 11),由于它的尺寸很小,增加了模型预测难度,导致其 R^2 小于其他病害,仅为 0.88。其他病害的 R^2 均不小于 0.97,整体病害预测 R^2 达到 0.98,表明本文所提出的实时语义分割算法,可应用于实际的草莓病害程度诊断任务中。

受野。此外,IBFFNet 采用交互式特征融合的方式实现上下文分支和空间分支间的特征互补,在一定程度上解决了特征提取过程中上下文分支细节特征不足和空间分支语义信息不强的问题。在单张 NVIDIA GTX1050 GPU 上,IBFFNet1 在 ImageNet 上以速度 189.4 f/s 获得 72.6% 的分类精度,超过多个先进的轻量化分类模型,说明了

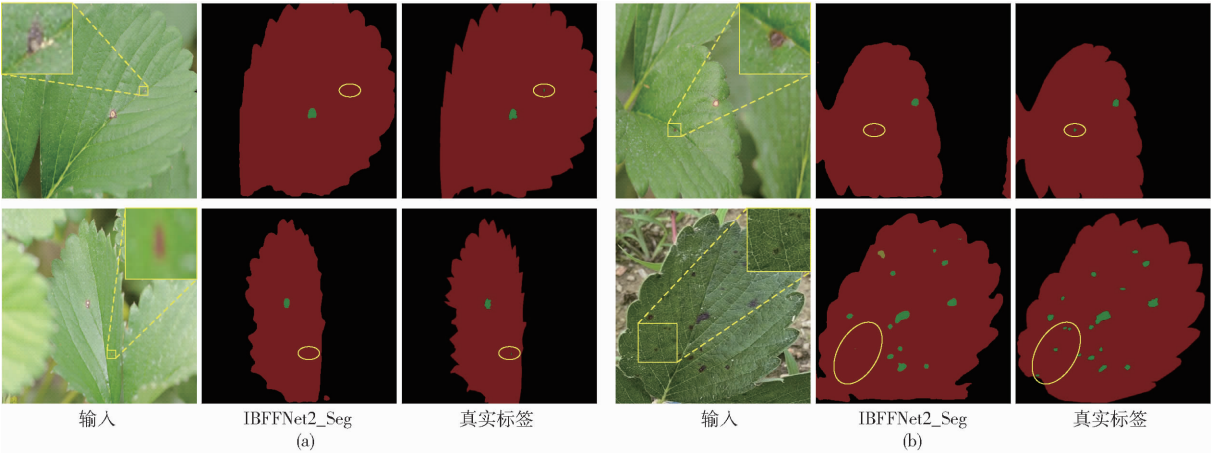


图 11 蛇眼病分割示例图

Fig. 11 Examples of white leaf spot segmentation results

IBFFNet 的高效性。

(2)以 IBFFNet 为骨干网络,本文构建草莓病害程度诊断模型 IBFFNet_Seg。针对叶片与背景对比度小、难以区分问题,提出一种边缘增强模块强化叶片和病灶边缘特征,丰富了空间分支中的边缘细节特征。针对部分病斑尺度小、形态多变问题,构建一个注意力简化的金字塔池化模块,提升了模型对小病斑特征的提取能力。在单张 NVIDIA GTX1050

GPU 上,IBFFNet2_Seg 能够以 77.8% 的 mIoU 每秒处理 40.6 幅图像,超过多个实时语义分割算法,满足实时病害程度诊断的需求。

(3)采用病斑与叶片区域的面积比作为病害程度值。IBFFNet2_Seg 的病害程度预测值与真实值的决定系数达 0.98,表明 IBFFNet2_Seg 的预测病害程度与真实程度相一致,可实现快速、准确的草莓病害严重程度诊断。

参 考 文 献

[1] 舒锐,焦健,臧传江,等. 我国草莓产业现状及发展建议[J]. 中国果菜, 2019, 39(1): 51 – 53.
SHU Rui,JIAO Jian, ZANG Chuanjiang, et al. The current situation and development suggestions of strawberry industry in China[J]. China Fruit and Vegetable, 2019, 39(1): 51 – 53. (in Chinese)

[2] IQBAL M, JAMSHAD M, ZAHID M A, et al. Biological control of strawberry crown rot, root rot and grey mould by the beneficial fungus *Aureobasidium pullulans*[J]. BioControl, 2021, 66(4): 535 – 545.

[3] 杜海顺,张春海,安文昊,等. 基于多层信息融合和显著性特征增强的农作物病害识别[J]. 农业机械学报, 2023, 54(7): 214 – 222.
DU Haishun,ZHANG Chunhai,AN Wenhao,et al. Crop disease recognition based on multi-layer information fusion and saliency feature enhancement[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(7):214 – 222. (in Chinese)

[4] 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)

[5] 杜甜甜,南新元,黄家兴,等. 改进 RegNet 识别多种农作物病害受害程度[J]. 农业工程学报, 2022, 38(15): 150 – 158.
DU Tiantian, NAN Xinyuan, HUANG Jiaying, et al. Identifying the damage degree of various crop diseases using an improved RegNet[J]. Transactions of the CSAE, 2022, 38(15): 150 – 158. (in Chinese)

[6] 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)

[7] 于明,李若曦,阎刚,等. 基于颜色掩膜网络和自注意力机制的叶片病害识别方法[J]. 农业机械学报, 2022, 53(8): 337 – 344.
YU Ming,LI Ruoxi,YAN Gang,et al. Crop diseases recognition method via fusion color mask and self-attention mechanism[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(8):337 – 344. (in Chinese)

[8] 任守纲,贾馥玮,顾兴健,等. 反卷积引导的番茄叶部病害识别及病斑分割模型[J]. 农业工程学报, 2020, 36(12): 186 – 195.
REN Shougang, JIA Fuwei, GU Xingjian, et al. Recognition and segmentation model of tomato leaf diseases based on deconvolution-guiding[J]. Transactions of the CSAE, 2020, 36(12): 186 – 195. (in Chinese)

[9] 雷雨,韩德俊,曾庆东,等. 基于高光谱成像技术的小麦条锈病病害程度分级方法[J]. 农业机械学报, 2018, 49(5): 226 – 232.
LEI Yu,HAN Dejun,ZENG Qingdong,et al. Grading method of disease severity of wheat stripe rust based on hyperspectral

- imaging technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(5): 226 – 232. (in Chinese)
- [10] 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.
- [11] GONÇALVES 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.
- [12] 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.
- [13] 刘斌, 徐皓玮, 李承泽, 等. 基于快照集成卷积神经网络的苹果叶部病害程度识别[J]. 农业机械学报, 2022, 53(6): 286 – 294.
- LIU Bin, XU Haowei, LI Chengze, et al. Apple leaf disease identification method based on snapshot ensemble CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(6): 286 – 294. (in Chinese)
- [14] RONNEBERGER O, FISCHER P, BROX T. U – Net: convolutional networks for biomedical image segmentation[C] // Proceedings of the Medical Image Computing and Computer-Assisted Intervention, 2015.
- [15] 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.
- [16] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C] // Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6230 – 6239.
- [17] LIN T, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C] // Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 936 – 944.
- [18] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C] // Proceedings of the Computer Vision, 2018: 833 – 851.
- [19] YU C, WANG J, PENG C, et al. BiSeNet: bilateral segmentation network for real-time semantic segmentation[C] // Proceedings of the Computer Vision, 2018: 334 – 349.
- [20] YU C, GAO C, WANG J, et al. BiSeNet V2: bilateral network with guided aggregation for real-time semantic segmentation[J]. International Journal of Computer Vision, 2021, 129(11): 3051 – 3068.
- [21] FAN M, LAI S, HUANG J, et al. Rethinking BiSeNet for real-time semantic segmentation[C] // Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 9716 – 9725.
- [22] PENG J, LIU Y, TANG S, et al. PP-LiteSeg: a superior real-time semantic segmentation model[J]. ArXiv:2204.02681, 2022.
- [23] PAN H, HONG Y, SUN W, et al. Deep dual-resolution networks for real-time and accurate semantic segmentation of traffic scenes[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(3): 3448 – 3460.
- [24] HUANG S, LU Z, CHENG R, et al. FaPN: feature-aligned pyramid network for dense image prediction[C] // Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision, 2021: 844 – 853.
- [25] TAKIKAWA T, ACUNA D, JAMPANI V, et al. Gated-SCNN: gated shape CNNs for semantic segmentation[C] // Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision, 2019: 5228 – 5237.
- [26] DENG J, DONG W, SOCHER R, et al. ImageNet: a large-scale hierarchical image database[C] // Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009: 248 – 255.
-

(上接第 216 页)

- [28] HUANG L, WU K, HUANG W, et al. Detection of fusarium head blight in wheat ears using continuous wavelet analysis and PSO – SVM[J]. Agriculture, 2021, 11(10): 998.
- [29] 范九伦, 张雪锋. 分段 Logistic 混沌映射及其性能分析[J]. 电子学报, 2009, 37(4): 720 – 725.
- FAN Jiulun, ZHANG Xuefeng. Piecewise Logistic chaotic map and its performance analysis[J]. Acta Electronica Sinica, 2009, 37(4): 720 – 725. (in Chinese)
- [30] TAVAZOEI M S, HAERI M. An optimization algorithm based on chaotic behavior and fractal nature[J]. Journal of Computational and Applied Mathematics, 2007, 206(2): 1070 – 1081.
- [31] ZHAO C, ZHAO H, WANG G, et al. Improvement SVM classification performance of hyperspectral image using chaotic sequences in artificial bee colony[J]. IEEE Access, 2020, 8: 73947 – 73956.
- [32] ZHANG C, DING S. A stochastic configuration network based on chaotic sparrow search algorithm[J]. Knowledge-Based Systems, 2021, 220: 106924.
- [33] 吴曙雯, 王人潮, 陈晓斌, 等. 稻叶瘟对水稻光谱特性的影响研究[J]. 上海交通大学学报(农业科学版), 2002, 20(1): 73 – 76.
- WU Shuwen, WANG Renchao, CHEN Xiaobin, et al. Effects of rice leaf blast on spectrum reflectance of rice[J]. Journal of Shanghai Jiaotong University (Agricultural Science), 2002, 20(1): 73 – 76. (in Chinese)