arXiv:2601.05482v1 [cs.CV] 9 Jan 2026

用于植物根系检测与分析的多图像超分辨率框架

Shubham Agarwal，*IEEE* 学生会员，Ofek Nourian、Michael Sidorov、Sharon Chemweno、Ofer Hadar，*IEEE* 高级会员，Naftali Lazarovitch 和 Jhonathan E. Ephrath

摘要—**深入理解植物根系对于推进土壤–植物相互作用、养分吸收及整体植物健康研究至关重要。然而，由于遮挡、土壤水分变化以及固有的低对比度等不利条件，地下环境中的根系精准成像仍面临持续挑战，这些因素限制了传统基于视觉的方法的有效性。本文提出一种新型地下成像系统，可捕捉植物根系的多个重叠视图，并集成一种基于深度学习的多图像超分辨率（MISR）框架，以增强根系的可见性并呈现更多细节。为训练和评估该方法，我们构建了一个合成数据集，用于模拟逼真的地下成像场景，并纳入影响图像质量的关键环境因素。所提出的 MISR 算法利用不同视图间的空间冗余，重建出结构保真度和视觉清晰度更高的高分辨率图像。定量评估表明，该方法优于最先进水平（SOTA）的超分辨率基线方法，在 CLIP-IQA 保持不变的情况下，BRISQUE 降低了 2.3%，表明图像质量有所提升，从而能够更好地对根系进行表型分析。这进一步有助于准确估算关键根系性状，包括根毛数量和根毛密度。所提出的框架为农业和生态研究中的稳健自动化地下植物根系成像与性状量化提供了有前景的方向。**

索引词—**多图像超分辨率、植物根系分析、目标检测、图像增强、地下摄像系统、模拟植物根系数据集**

一、引言

在资源有限、土壤质量不断下降且粮食需求日益增加的背景下，农业中的人工智能（AI）正逐渐成为一个重要的研究领域。目前已经出现了多种基于深度学习的算法，用于分析植物特征，例如识别叶部病害、虫害、果实品质等。利用无人机获取的航空影像也被用于监测森林和农田区域 [1], [2]。这些方法极大地拓展了我们对植物生长的认识，从而提高产量。相较之下，地下植物根系的分析是一个探索较少的领域。根系对植物的整体健康状况起着至关重要的作用 [3], [4], [5], [6]。除了提供锚定作用外，根系还会从土壤中吸收水分和养分。根系还具有根毛，有助于更高效地吸收养分。由于植物根系位于地下，对其进行分析较为困难，而成像系统还面临潮湿土壤、遮挡和低对比度等诸多问题 [7], [8]。

在本研究中，我们探讨了将多图像超分辨率（MISR）技术应用于提升地下 RGB 图像质量，以便进行植物根系分析。为此，我们使用了一套名为 Multi-ImageRootCam 的硬件系统，专为采集地下环境中植物根系的高分辨率、重叠 RGB 图像而设计。Multi-Image RootCam 装置能够采集根系结构在空间上相互重叠的多个视角，这些图像将作为多图像超分辨率流程的输入。

除了真实图像采集外，我们还构建了一个合成数据集，用于模拟 Multi-Image RootCam 生成的重叠地下根部图像。该模拟数据集用于训练我们提出的多图像超分辨率模型，从而支持受控实验和可扩展的模型开发。我们的多图像超分辨率架构以最先进水平的深度残差通道注意力变换器（DRCT）模型为基础，该模型最初是为单图像超分辨率任务开发的。我们扩展并调整了这一架构，使其能够处理多个重叠图像输入，利用不同视角之间的互补信息，相较于单图像方法更有效地提升分辨率。通过大量实验，我们证明，与现有超分辨率技术相比，我们的模型在定量指标和视觉保真度方面均显著领先。

本研究的主要贡献总结如下：

- 我们设计了一种新型 RootCam成像系统，可采集重叠的地下植物根系 RGB 视图，用于多图像超分辨率。
- 我们构建了一个合成重叠根系数据集，用于模拟地下成像条件，以支持稳健的模型训练。
- 我们提出了 MI-DRCT框架，将多图像融合与对齐模块集成到 DRCT骨干网络中，以利用重叠视图之间的亚像素位移。
- 所提出的方法在定量指标和感知效果方面均取得了显著提升，增强了根系结构的清晰度，并提高了下游分割准确率。

我们的框架在提升根部图像质量方面取得了显著成果，在

稿件收稿日期：2025年6月30日；修订日期：2025年10月20日和2025年11月16日；通讯作者：*Shubham Agarwal*。Shubham Agarwal、Ofek Nourian、Michael Sidorov 和 Ofer Hadar 就职于以色列贝尔谢巴84105号内盖夫本-古里安大学电气与计算机工程学院（电子邮箱：agarwals@post.bgu.ac.il、ofekno@post.bgu.ac.il、sidorov@post.bgu.ac.il、hadar@bgu.ac.il）。Sharon Chemweno、Naftali Lazarovitch 和 Jhonathan E. Ephrath 就职于以色列塞代博克84105号内盖夫本-古里安大学雅各布·布劳斯坦沙漠研究所法国农业与旱地生物技术协会研究所（电子邮箱：chemweno@post.bgu.ac.il；lazarovi@bgu.ac.il；yoni@bgu.ac.il）。

0000–0000/00$00.00 © 2025年 IEEE

BRISQUE [9]评分上超越了最接近的方法 2.3%，且 CLIP-IQA [10] 评分相同。结果表明，我们的方法应用于地下根系图像分析时具有优越性。

我们的系统所提供的增强图像质量能够更好地量化关键的根表型性状，例如根毛数量和表面积。通过将多图像 RootCam 硬件与改进的多图像超分辨率流程相结合，我们实现了对地下植物根系的详细可视化，这对于推动根系表型研究具有巨大潜力。

II. 相关工作

A. 超分辨率

经典的多图像超分辨率（MISR）方法可追溯至 20 世纪 70 年代 [11], ，当时通过融合多幅低分辨率（LR）图像来校正混叠伪影。尽管这些早期方法在小规模改进方面颇为有效，但它们严重依赖手工设计先验，并且对噪声和运动不一致缺乏稳健性。

随着 Transformer 的出现 [12], ，其在捕获全局依赖关系方面的成功促使其迅速应用于视觉任务 [13], [14]。因此，许多研究探索了用于超分辨率（SR）的基于 Transformer 的架构 [15]–[16], [17], [18], [19], [20]。然而，从基于 CNN 的局部建模过渡到基于 Transformer 的全局推理仍并非易事，在有限数据上训练时往往会导致较高的计算成本和不稳定收敛。

Yang 等人 [21]提出了一种纹理 Transformer，通过将纹理特征编码为 Transformer 的键、值和查询，缓解超分辨率中的纹理退化。尽管这种方法能有效恢复精细细节，但它严重依赖参考图像，在缺少参考图像或参考图像不一致时效果不佳。类似地，Chen 等人 [22] 提出了图像处理 Transformer（IPT），这是一种面向低层级任务的大规模预训练模型。尽管 IPT 展现出令人瞩目的结果，但其庞大的规模（1.15 亿个参数）以及对大型多任务数据集的依赖，限制了其实用性，也阻碍了它在资源受限环境中的部署。

对于视频超分辨率，Cao 等人 [23]通过 VSR Transformer 将基于 Transformer 的推理扩展到时空数据。然而，他们依赖卷积神经网络提取的特征，限制了模型捕捉长期时间依赖关系的能力；分块注意力机制也导致局部上下文碎片化。Liang 等人 [24] 将 Swin Transformer[14] 应用于超分辨率，结合卷积神经网络的浅层特征与基于 Transformer 的深层表征。尽管这种混合结构提升了局部特征转换能力，但由于采用分层窗口移位，计算效率仍然不理想。

Uformer 模型 [25] 采用带有 LeWin 模块的 U 形 Transformer 拓扑结构，以平衡局部与全局特征学习。其局部自注意力机制降低了复杂度，但分层结构往往会导致精细纹理区域出现过度平滑。总体而言，现有基于 Transformer 的超分辨率方法揭示了全局上下文建模与空间细节保留之间的权衡，而 DRCT 框架通过引入贯穿 Transformer 层的高效密集残差连接，直接应对这一挑战，以确保多图像一致性。

B. DRCT

我们以密集残差连接 Transformer（DRCT） [26], 为基础，该模型最初专为单图像超分辨率而设计。DRCT 通过在各层之间引入密集残差连接，缓解了深层 Transformer 网络中普遍存在的信息瓶颈，从而改善梯度流并增强多层特征复用。这使模型能够有效融合局部与全局上下文，同时不损失空间精度。

尽管 DRCT 在单图像超分辨率领域取得了最先进的成果，但尚未扩展至多图像领域。我们提出的 MI-DRCT 通过引入图像间特征对齐与融合弥补了这一不足，使模型能够利用多个低分辨率输入中的互补信息。这一设计不仅增强了细粒度纹理恢复能力，还提升了模型在地下根系成像场景中应对噪声和光照变化的鲁棒性。

C. 根系图像分析

传统上，根系成像主要采用经典图像分析方法 [27], [28], [29], [30]。早期方法依赖形态学运算或几何拟合（例如基于圆柱体或边缘的方法）来推断直径、面积和长度等结构属性。尽管这些方法对孤立样本有效，但面对复杂根系时，由于遮挡、噪声以及根与土壤之间对比度较低，往往难以奏效。

探地雷达（GPR）技术 [31], [32], [33]–[34], [35], [36] 能够实现无损地下成像，但其性能对土壤纹理、水分和根系密度高度敏感，限制了其泛化能力。这些依赖因素使得在不同环境条件下进行一致的定量分析变得困难。

深度学习的最新进展重新推动了基于 RGB 的地下成像 [37], [38], ，利用超分辨率（SR）等增强模型来恢复结构细节。然而，大多数现有方法仍处于初步阶段 [39], [40], 且难以应对域变异，尤其是土壤异质性和水分干扰所造成的影响。数据采集也缓慢且不一致，使得稳健的模型训练颇具挑战。我们的 MI-DRCT框架通过多图像融合提升分辨率和对比度，弥补了这些不足，从而能够在不利成像条件下更可靠地进行特征提取。

Li 等人 [41]观察到，轻量级超分辨率模型通常会出现严重的精度下降，而大容量模型则会产生高昂得难以接受的计算成本。他们提出的多尺度特征选择网络（MFSN）实现了较好的权衡，但仍依赖单图像输入，因此在土壤或光照条件变化时的鲁棒性受到限制。类似地，Wu 等人 [42] 提出了多尺度非局部注意力网络（MSNLAN），用于捕获长距离像素依赖关系，但其架构引入了跨尺度冗余，增加了推理延迟。这些研究共同凸显了对以下架构的需求：在保持空间



图 1。部署于农业田地中的 RootCam。请注意，有三个可见的 RootCam 设置，每个设置均以绿色矩形标出。

保真度的同时兼具计算效率——这正是我们提出 MI-DRCT 所秉持的核心原则。

III. 方法

我们的方法由四个不同部分组成，分别是用于采集真实植物根系图像的 Multi-Image RootCam、用于训练的合成根系图像生成、用于图像增强的 MI-DRCT，以及用于获取根系特征的根系图像分析。以下小节将分别对这些部分进行详细介绍。

A. Multi-Image RootCam

如图 1 所示，Multi-Image RootCam（CrystalVision，以色列萨马尔）是 RootCam [30] 的扩展设备。它由一根底部封闭的透明管组成，高 100 cm，外径为 60 mm，内径为 54 mm。该管垂直安装于土壤中。位于土壤上方的管体顶部内侧涂成黑色，以防止光线进入；外侧涂成白色，以反射多余光线。随后，将 RootCam 设备小心地插入管内进行成像（图 1）。

RootCam 设备本身包括一个旋转滑架，滑架搭载一台高分辨率摄像机（2592 × 1944像素，空间分辨率为 0.01 mm，2500 dpi）和集成式 LED 光源（白光和紫外光）。为避免干扰植物生长，仅在图像采集期间开启照明。摄像机沿管道移动时，每隔 20 mm 采集一次图像。摄像机的位置和运动（垂直及旋转）由 RootCam 软件界面控制，该界面还提供图像重叠量、超分辨率和实时预览等功能。

所有采集的图像都存储在内置的 Raspberry Pi 计算机中，可通过局域网在本地访问，也可通过 Wi-Fi、VNC 或 Raspberry Pi Connect 远程访问。

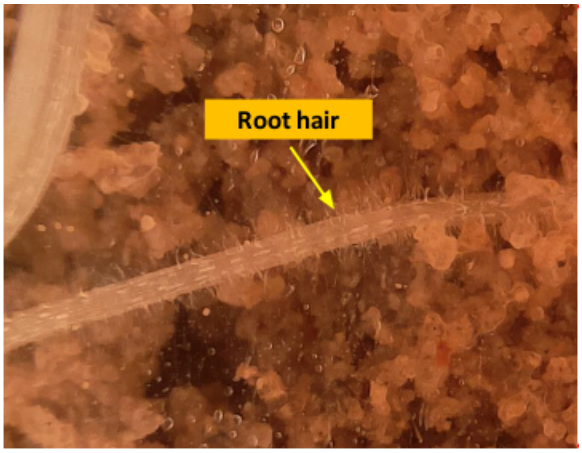

图 2. RootCam 拍摄的根部图像示例，展示了根毛。

系统还支持自动上传至 Dropbox 等云存储服务。每个图像文件都包含元数据，包括管内深度、旋转步进量和采集时间。该设备使用 12V 直流电源（5A）供电，运动控制功能可使相机在管内实现亚像素定位。这套完整的系统能够在不同深度对根系动态进行连续、无损监测，并实现精确的空间和时间分辨率。

B. 合成根系图像生成

尽管我们可以使用 RootCam 系统，但要获取足够多的真实图像来训练深度神经网络（通常需要数千张），仍颇具挑战。根的生长速度缓慢，加上土壤和光照条件需要多样化，因此难以大规模采集真实数据。为克服这一限制，我们模拟地下环境并生成合成根部图像。具体而言，我们使用 RootCam 系统采集各种真实背景样本，并运用统计模型叠加不同形态和根毛密度的根。这些合成图像大幅扩充了训练集，并改善了模型收敛。

我们首先生成一张超高分辨率的合成根部图像。然后，通过对该图像进行裁剪并施加受控的亚像素垂直位移，生成多张相互重叠的根系图像。此过程包括以下步骤：

- 图像裁剪：裁剪合成根部图像中以根系结构为中心的一个子区域，以聚焦于感兴趣的面积。
- 对图像施加奇数个像素的垂直位移。使用奇数个像素进行位移，可确保随后将图像缩小 2 倍时产生 0.5 的亚像素位移。这会为位移后的图像带来更多独特变化，有利于超分辨率任务。
- 下采样与插值：使用 cv2.INTER_AREA 插值对平移后的裁剪图像进行下采样。该方法会计算下采样图像中每个新像素对应区域的平均像素值。这有助于最大限度地减少混叠伪影并保留图像细节

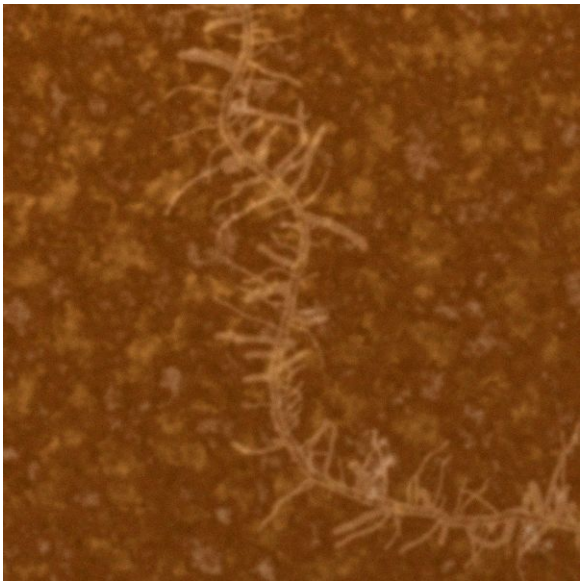

图 3. 由我们的合成根系生成算法生成的图像。

在下采样过程中，这对于保留亚像素位移信息至关重要。

使用合成生成数据进行训练，不仅弥补了真实样本的匮乏，还增强了 MI-DRCT 模型的鲁棒性和泛化能力。通过模拟多种光照变化、土壤纹理和根系形态，合成数据集使模型接触到多样的视觉条件，而这些条件在有限的真实数据集中可能无法得到充分体现。这种多样性使模型能够学习到不变特征，从而能够抵御真实 RootCam 图像中的光照变化、遮挡和噪声。因此，在合成数据上预训练的模型，在针对真实世界的根系成像场景进行微调或直接应用时，表现出更强的稳定性和适应性。

C. MI-DRCT

如前文所述，RootCam 可用于采集植物根系的多幅重叠图像，我们也可以生成能够模拟这一场景的合成图像。现在，我们设计了一种算法，通过利用重叠图像中存在的冗余信息和亚像素信息来提升图像质量。我们以 DRCT [26] 为基础——它是目前单图像超分辨率领域最先进水平的框架——并对其进行改造，使其能够利用地下根系数据中的冗余信息和亚像素信息。新框架称为多图像深度残差通道变换（MI-DRCT）。MI-DRCT 的架构如图 4 所示。

如图 4 所示，MI-DRCT框架包含两个主要阶段。这些阶段旨在利用重叠输入图像之间的空间冗余，增强特征表示，并最终提升超分辨率性能。

1) 浅层特征提取与对齐：第一阶段如图 4(a) 所示，涉及提取和对多张相互重叠的低分辨率（LR）输入图像中的浅层特征进行对齐。每张输入图像均独立通过一个共享卷积层，以提取初始浅层特征。该共享层确保整个图像序列中的特征表示保持一致。

获得初始特征后，将输入序列中居中（或索引居中）的图像所对应的特征图指定为参考特征图。其余 N–1 个特征图通过特征对齐模块与该参考特征图对齐。该模块旨在处理由相机运动或视差引起的图像间小幅空间错位。

对齐过程利用相位相关模块，在傅里叶域中估计查询特征图与参考特征图之间的相对位移。鉴于输入的特性（例如根系图像），我们观察到错位主要沿垂直轴发生。因此，对齐时仅考虑估计位移的垂直分量。为提高对齐精度，尤其是在存在亚像素位移的情况下，相位相关处理前会将查询特征图和参考特征图均按 2 倍因子上采样。

对齐后，所有特征图沿通道维度进行拼接。对于 N 张图像，每张图像的特征维度为 f，所得拼接张量的维度为 N×f。随后，将该拼接特征图输入第二个卷积层，将通道维度重新降至 f，从而得到最终浅层特征。

2) 深层特征提取与图像重建：第二阶段如图4(b)所示，包括更深层的特征提取和图像重建。最终浅层特征经过一系列残差深层特征提取组（RDGs）处理，该结构最初由 DRCT [26] 框架提出。这些 RDGs 旨在提取复杂的层次特征，从而有助于高分辨率（HR）输出中的细粒度细节恢复。

为保留低层次信息并促进梯度流，采用跳跃连接将浅层特征与深层特征进行拼接。随后，将这一融合后的特征表示输入图像重建模块，生成与输入序列中中心图像对应的最终超分辨率输出。

需要注意的是，该架构中的所有卷积层均采用 3×3 卷积核，并额外进行零填充。这种配置可确保在每个卷积阶段，输出张量的空间分辨率（即高度和宽度）与输入保持一致。

D. 根系图像分析

为分析地下植物根系的形态，我们开发并实现了一种基于深度学习的实例分割模型。该模型使用专门生成的合成数据集进行训练，以模拟地下根系结构，包括主根和细小的根毛。给定一幅地下根系的输入图像，模型会生成逐像素实例分割掩码，清晰地识别根体及其根毛。

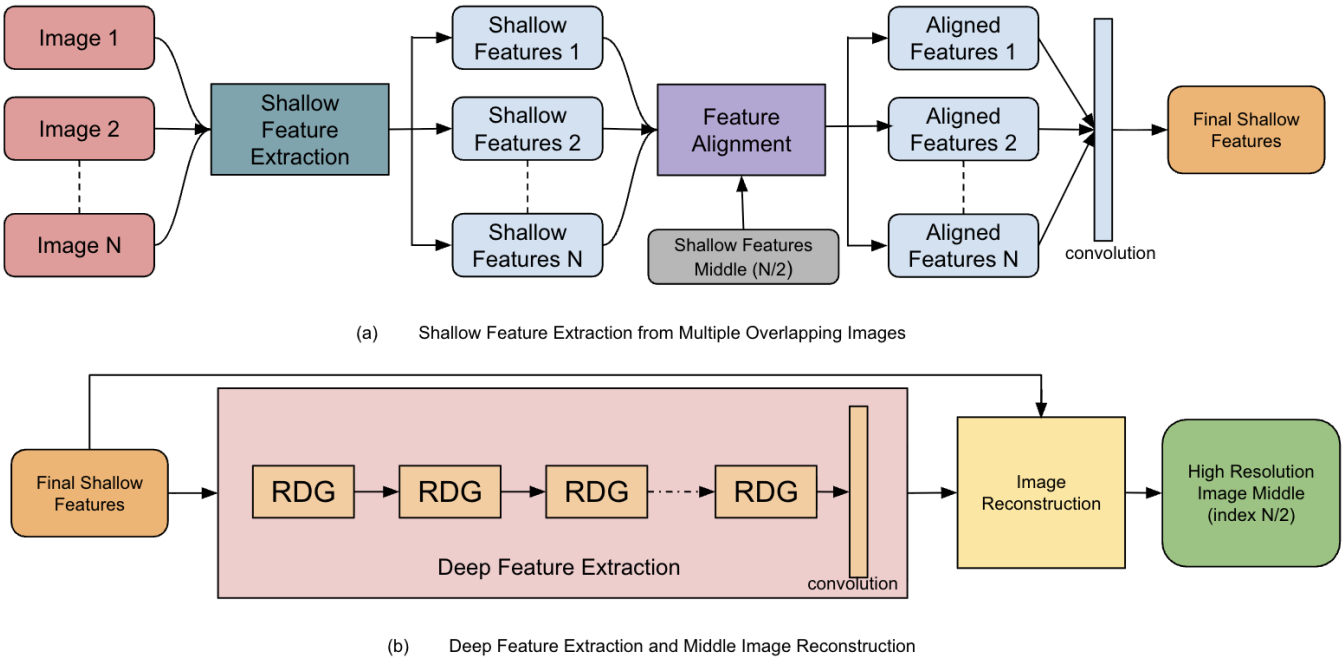

图4. MI-DRCT 的架构。我们将其分为两个阶段。(a) 步骤进行浅层特征提取与对齐移位特征。(b) 步骤以组合后的浅层特征为输入，并使用残差密集组（RDG）模块提取深层特征。随后将深层特征与浅层特征组合，生成超分辨率图像。

分割完成后，模型会计算根毛的平均面积（以像素为单位）。随后，根据图像已知的空间分辨率，将这一基于像素的测量值转换为物理单位（平方毫米）。所得指标可定量估算根毛密度，并可作为评估整体植物健康状况和养分吸收能力的有用指标。

图6展示了以真实根部图像为输入时，MI-DRCT及其他算法对一个样本的分割结果。如图所示，主要根系结构以浅蓝色勾勒，根毛则以深蓝色清晰标出，从而能够直观地区分这两种解剖结构组成部分。

IV. 实验设置

所有实验均在配备 NVIDIA RTX 3090 GPU 的 Windows 工作站上进行。为训练我们提出的 MI-DRCT 模型，我们构建了一个由 10,000 个根部图像三元组组成的合成数据集用于训练，另有 2,000 个三元组用于验证。每个三元组由 3 张图像组成，MI-DRCT 中的垂直位移为 $N = 3$（图4）。重叠图像数量设为 3，以最大限度增加用于评估的独立样本数量。这种合成设置可实现受控实验，并确保各项评估的一致性。在合成数据上完成这些实验后，我们将模型应用于一个真实数据集，该数据集包含甜椒植株的 50 个地下根部图像三元组。该实验验证了系统在真实环境中的有效性。

MI-DRCT 模型训练了 15 个训练轮次，并选取在验证集上表现最佳的模型检查点用于最终推理。在训练过程中，我们采用批量大小为 8，并使用 Adam 优化器 [43] 对网络进行优化。初始学习率设为 1×10^{-3}，且未采用学习率衰减。

每幅输入图像均裁剪为空间分辨率为64×64 像素的图像，并训练模型生成尺寸为 128×128的对应高分辨率输出，对应的放大因子为 2。得益于其全卷积架构，MI-DRCT 模型与分辨率无关，在推理过程中无需修改架构即可无缝适配任意尺寸的输入。

为进行对比评估，我们采用了基线 DRCT 模型 [26]，并使用其默认超参数设置。具体而言，嵌入维度设为 96，窗口大小设为 16，Transformer 配置保持原始设定，深度和注意力头数量均设为 (6,6,6,6)。放大因子同样固定为 2，以确保与我们提出的模型输出分辨率保持一致。

V. 结果

如前文所述，我们使用合成数据训练模型。我们在留出的合成数据上测试多图像超分辨率模型，并将其与其他最先进的单图像和多图像超分辨率算法进行比较。我们还使用 RootCam 拍摄的真实根系图像对这些模型进行测试和比较。随后，我们在增强图像上检测并分析根系特征。在我们的实验中，CLIP-IQA 指标将图像的嵌入表示与两个固定的文本提示“好照片”和“坏照片”进行比较。它计算图像与这两个提示之间的余弦相似度，并输出图像更接近正向提示的概率，以此表示图像的感知质量分数。

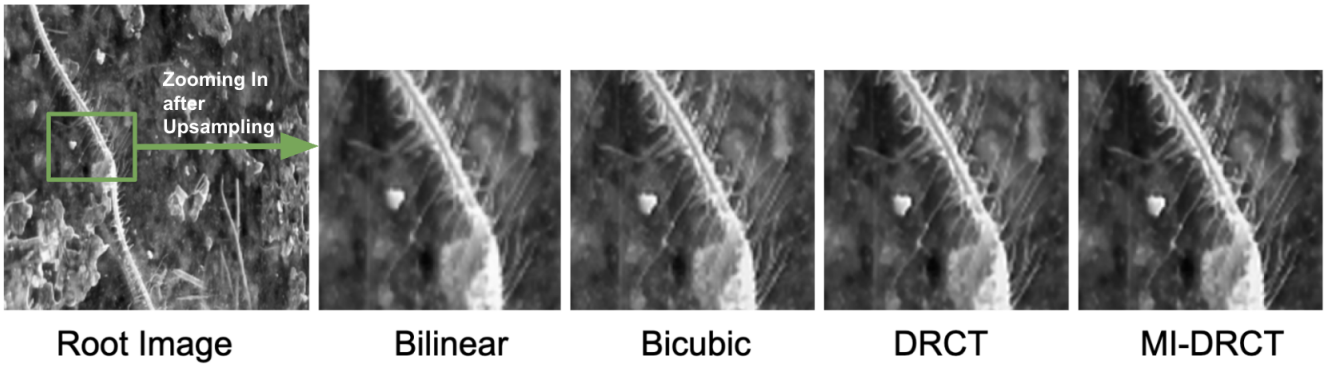

图 5。 不同方法的定性比较 超分辨率技术在 RootCam 拍摄的甜椒植株真实图像上的定性比较。我们可以看到，MI-DRCT 框中的对比度更高 r MI-DRCT 框中的对比度，其中 ，从而增强了根毛。DRCT 的对比度也不错，但质量略低于 MI-DRCT。

该图像越接近正向提示，其感知质量评分就越高。

A. 合成根系数据集上的结果

我们使用有参考（均方误差、峰值信噪比和结构相似性指数）以及无参考（BRISQUE 和 CLIP-IQA）图像质量指标，在单图像和多图像超分辨率任务上，将新提出的 MI-DRCT 算法与现有方法进行了比较。

表I 在有参考和无参考图像质量指标上比较所提出的 MI-DRCT 框架与其他方法（合成测试数据集）。

框架	均方误差 ↓	峰值信噪比	结构相似性指数	BRISQUE ↓	CLIP-IQA
双线性	25.28	36.80	0.93	46.92	0.32
双三次	19.84	38.05	0.94	43.79	0.36
DRCT	18.14	38.22	**0.95**	44.23	0.38
MI-DRCT	**16.31**	**38.96**	**0.95**	**43.13**	**0.39**

在表I中，我们比较了我们的框架（MI-DRCT）与现有方法的超分辨率质量。可以清楚地看到，在无参考和有参考图像质量指标方面，MI-DRCT框架均优于所有现有方法。在有参考质量指标方面，MI-DRCT的均方误差（MSE）更低，峰值信噪比（PSNR）和结构相似性指数（SSIM）更高，优于当前最先进水平的DRCT [26]及其他基线方法。对于BRISQUE [9]和CLIP-IQA [10], 这类无参考图像质量指标，也观察到了类似的表现，MI-DRCT取得了最佳结果。

值得注意的是，DRCT方法是当前单图像超分辨率的最先进基准，而利用MI-DRCT通过多幅具有亚像素位移的图像来增强信息，则展示了一种有前景的方向，可将其他单图像超分辨率框架拓展到多图像场景。

B. 真实根系数据集结果

我们还将使用合成数据训练的 MI-DRCT框架的结果与其他方法在真实地下根系测试图像上的结果进行了比较，这些图像由 Multi-Image RootCam 获取。真实图像来自一株甜椒植株。在此，我们只能采用无参考图像质量指标，因为超分辨率评估没有参考图像。因此，我们的比较仅限于 BRISQUE 和 CLIP-IQA 分数。结果如表II所示。

表II 展示了所提出的 MI-DRCT 超分辨率框架与其他方法在真实图像测试数据集上的无参考图像质量指标对比。

框架	BRISQUE ↓	CLIP-IQA
双线性	46.48	0.30
双三次	44.61	0.31
SwinIR [24]	46.03	0.33
BSRGAN [44]	46.57	0.36
DRCT [26]	45.33	**0.38**
MI-DRCT	**44.50**	**0.38**

我们可以清楚地看到，MI-DRCT算法在真实数据上的表现同样最佳。其BRISQUE值最低，为44.50，CLIP-IQA得分为0.38。原始DRCT算法的CLIP-IQA得分相同，但BRISQUE值更高。图5展示了多种超分辨率算法生成的原始图像和增强图像的定性比较。与双线性和双三次方法相比，MI-DRCT明显改善了根毛可视化效果。尽管DRCT和MI-DRCT在视觉上看起来相似，但定量指标证实了MI-DRCT的优越性能，突出了使用多张移位图像的优势。

值得注意的是，合成测试数据集与真实测试数据集的表现十分接近，且 BRISQUE 和 CLIP-IQA 得分相似。由于真实数据难以获取，这一结果表明合成生成的根系数据同样适用于训练自动根毛检测模型，因此具有重要意义。

C. 植物根系的检测与分析

所提出的 MI-DRCT算法显著增强了分割模型准确检测精细结构的能力，尤其是根毛。这些结构通常由于尺寸较小且在原始图像中的对比度较低而难以辨别。如图6所示，该分割

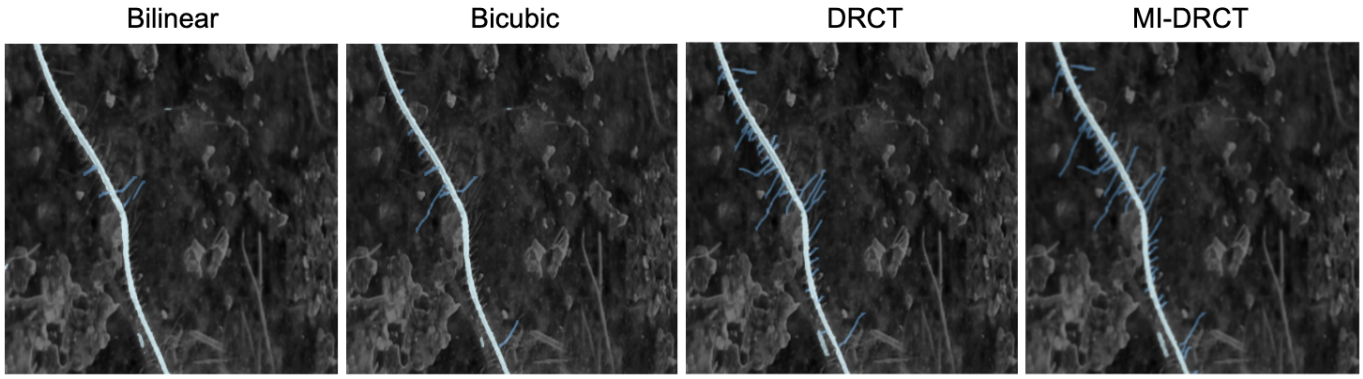

图6. 不同超分辨率方法下图像分割算法性能的比较。MI-DRCT 增强后的图像能够更准确地自动检测根和根毛。

表 III 图6中所识别根毛特征表。

根系性状	双线性	双三次	DRCT	MI-DRCT
根数（**人工专家：1**）	1	1	1	1
根毛数量（**人工专家：77**）	4	8	38	**44**
根毛总长度（mm）	7.88	14.71	58.54	71.47
根毛平均长度（mm）	1.97	1.83	1.54	1.62
平均根毛面积（mm^2）	0.29	0.25	0.20	0.19

算法在采用双线性和双三次插值时只能检测到少量根毛。使用 DRCT 后，根毛检测效果大幅提升，共检测到 38 根根毛。相比之下，使用 MI-DRCT 时，在人工专家标记的 77 根根毛中检测到 44 根。由于成像系统的空间标定（单位为 mm）已知，因此可以将分割输出可靠地映射到物理单位。这使得自动提取根毛长度和面积等相关指标成为可能。详细分析见表 III。使用 MI-DRCT 时，根毛总长度为 71.47 mm，根毛平均长度为 1.62 mm。我们还对其他超分辨率算法进行了类似分析。根毛长度和面积过于精细，即使由专家进行分析，也会产生较大误差。我们的框架具备捕捉精细形态特征的能力，这对于研究《植物生理学》、养分吸收以及整体根系构型至关重要。

VI. 消融研究

A. 亚像素对齐的重要性

我们开展了消融实验，以确定亚像素移位模块（在浅层特征对齐之前进行 2 倍上采样）在 MI-DRCT 方法中的重要性。结果如表 IV 所示。

表 IV 展示在合成地下根系图像数据集上，MI-DRCT 超分辨率框架中考虑亚像素位移的重要性的表格。

框架	均方误差 ↓	峰值信噪比	结构相似性指数	BRISQUE ↓	CLIP-IQA
MI-DRCT（无亚像素）	17.65	38.45	0.95	43.80	0.37
MI-DRCT	16.31	38.96	0.95	43.13	0.39

表 IV 突出了在所提出的 MI-DRCT框架中考虑亚像素位移的关键作用，评估基于合成的地下根部图像数据集进行。我们比较了 MI-DRCT框架的两个变体：一种直接聚合多图像输入而不考虑亚像素偏移，另一种则在融合前对各帧之间进行亚像素对齐。

结果表明，纳入亚像素位移可显著提升整体图像重建质量。具体而言，均方误差（MSE）降低了 7.5%，说明逐像素重建误差有所减小。两个变体的结构相似性指数（SSIM）（0.95）和峰值信噪比（PSNR）（38.9 dB）均保持一致，表明其结构信息和信号信息表现相当。BRISQUE（衡量自然度的指标）从 43.80 降至 43.13（越低越好），CLIP-IQA 得分则从 0.37 升至 0.39，表明图像内容与语义意义的契合度更高。总体而言，具备亚像素感知能力的模型表现更优。

这些改进证实，多个低分辨率帧中的亚像素位移能够捕获非冗余且互补的空间细节，从而用于重建质量更高的超分辨率图像。若不补偿这些位移，聚合后的帧所提供的信息大多相互重叠，限制了多图像超分辨率的潜在收益。因此，亚像素对齐对于有效利用多图像输入中额外的帧间信息至关重要。

VII. 结论

本文提出了一种创新框架，通过利用多张相互重叠的 RGB 图像，提升植物根系和根毛的检测与分析能力。根和根毛的表型分析在植物生物学与农业研究中发挥着关键作用，但要获取地下根系的详细、准确图像，

仍是一项重大挑战，原因在于遮挡、对比度低以及土壤干扰。为解决这一问题，我们提出了一种多图像增强方法，融合多个 RGB 视图中的信息，以重建高质量的地下根系结构表示。该方法无需采用侵入式成像技术，即可更准确地定位和计数根毛。

我们框架的核心是 MI-DRCT（多图像深度残差通道变换）算法，这是一种基于深度学习的超分辨率模型，利用模拟复杂地下环境的合成数据集进行训练。使用合成生成的根系图像进行训练，使 MI-DRCT 模型能够接触到更丰富多样的纹理、光照条件和根系结构，从而增强其鲁棒性和泛化能力。这种多样性使模型能够学习到更多不变特征，提升其在真实 RootCam 数据上的性能和稳定性。通过大量实验，我们证明，MI-DRCT 在重建精细尺度根系结构方面优于现有最先进算法。该模型在合成环境和真实成像条件下均表现出卓越性能，能够生成高保真视觉结果，助力根系分割、特征提取和根毛定量等下游任务。

我们的结果表明，在结构化且可学习的框架中整合多张低分辨率 RGB 图像，能够显著提升根系图像的视觉保真度和分析实用性。借助多视图中蕴含的互补信息，所提出的 MI-DRCT 框架能够有效重建清晰度和细节更佳的高分辨率根系结构。这对于非侵入式根系成像尤为有益，因为保留精细结构特征对于后续分析至关重要。MI-DRCT 框架还有助于更准确、可靠地解读根系构型，这对于根长估算、分枝模式检测和生长监测等定量表型分析任务至关重要。此外，该框架采用模块化设计且易于扩展，非常适合部署于高通量植物表型分析流程中，以便高效、一致地处理大量根系图像。

总的来说，所提出的方法代表了根部成像技术的一项重大进步，为研究人员和农学家提供了一种实用且可靠的解决方案，帮助他们了解不同土壤和环境条件下的根系动态。

致谢

本项目获得农业与粮食安全部资助：开发用于优化灌溉和施肥的根系特征自动表征系统。本研究还部分获得以色列高等教育委员会规划与预算委员会（VATAT）的资助，资助通过以色列数字农业中心计划提供。

参考文献

[1] E. Singh、A. Pratap、U. Mehta 和 S. I. Azid，《采用图像处理和机器学习技术进行作物喷洒的智能农业无人机：实验验证》，Drones，第5卷，第2期，第13页，2024年。[2] B. Alsadik、F. J. Ells¨aßer、M. Awawdeh、A. Al-Rawabdeh、L. Alma- hasneh、S. Oude Elberink、D. Abuhamoor 和 Y. Al Asmar，《利用无人机开展病虫害监测的遥感技术：聚焦枣椰树的综述》，Remote Sensing，第16卷，第23期，第4371页，2024年。[3] T. Beeckman 和 A. Eshel 编，《植物根系：隐藏的另一半》，第5版。博卡拉顿：CRC出版社，2024年。[4] D. Zhao、P. de Voil、V. O. Sadras、J. A. Palta 和 D. Rodriguez，《根系性状的可塑性及其对作物产量和产量稳定性的影响》，Plant and Soil，第504卷，第1–2期，第1–16页，2025年。[5] B. El Amrani，《探究根系构型可塑性在植物适应环境约束中的重要性》，Plant Species Biology，第38卷，第5期，第234–244页，2023年。[6] E. L. Fry、A. L. Evans、C. J. Sturrock、J. M. Bullock 和 R. D. Bardgett，《根系构型决定植物对干旱响应的可塑性》，Plant and Soil，第433卷，第189–200页，2018年。[7] L. Li、Q. Zhang 和 D. Huang，《植物表型分析成像技术综述》，Sensors，第14卷，第11期，第20 078–20 111页，2014年。[8] A. Gillert、B. Peters、U. F. von Lukas、J. Kreyling 和 G. Blume-Werry，《在微根管图像中跟踪植物根系的生长和衰亡》，载于2023年IEEE/CVF冬季计算机视觉应用会议（WACV），2023年，第3688–3697页。[9] A. Mittal、A. K. Moorthy 和 A. C. Bovik，《盲/无参考图像空间质量评估器》，载于2011年第45届阿西洛马会议（信号、系统与计算机）会议录（ASILOMAR），2011年，第723–727页。[10] J. Wang、K. C. Chan 和 C. C. Loy，《探索CLIP在评估图像观感方面的应用》，丛书：AAAI’23/IAAI’23/EAAI’23。AAAI出版社，2023年。 [在线]。链接：https://doi.org/10.1609/aaai.v37i2. 25353[11] S. Peleg、D. Keren 和 L. Schweitzer，《利用亚像素运动提高图像分辨率》，Pattern recognition letters，第5卷，第3期，第223–226页，1987年。[12] A. Vaswani、N. Shazeer、N. Parmar、J. Uszkoreit、L. Jones、A. N. Gomez、Ł. Kaiser 和 I. Polosukhin，《注意力就是一切》，神经信息处理系统进展，第30卷，2017年。[13] A. Dosovitskiy、L. Beyer、A. Kolesnikov、D. Weissenborn、X. Zhai、T. Unterthiner、M. Dehghani、M. Minderer、G. Heigold、S. Gelly 等，《一张图像胜过千言万语：面向大规模图像识别的Transformer》，arXiv预印本 arXiv:2010.11929，2020年。[14] Z. Liu、Y. Lin、Y. Cao、H. Hu、Y. Wei、Z. Zhang、S. Lin 和 B. Guo，《Swin Transformer：采用移位窗口的分层视觉变换器》，载于IEEE/CVF国际计算机视觉会议论文集，2021年，第10 012–10 022页。[15] F. Yang、H. Yang、J. Fu、H. Lu 和 B. Guo，《学习用于图像超分辨率的纹理变换器网络》，载于IEEE/CVF计算机视觉与模式识别会议论文集，2020年，第5791–5800页。[16] J. Cao、J. Liang、K. Zhang、Y. Li、Y. Zhang、W. Wang 和 L. V. Gool，《采用可变形注意力变换器的基于参考图像超分辨率》，载于欧洲计算机视觉会议。Springer，2022年，第325–342页。[17] K. Cheng、L. Yu、Z. Tu、X. He、L. Chen、Y. Guo、M. Zhu、N. Wang、X. Gao 和 J. Hu，《用于图像超分辨率的有效扩散变换器架构》，2024年。[18] S. Jiang、S. Xu 和 X. Wang，《RBSformer：用于原始图像超分辨率的增强型Transformer网络》，载于2024年IEEE/CVF计算机视觉与模式识别会议研讨会（CVPRW），2024年，第6479–6488页。[19] Z. Lu、J. Li、H. Liu、C. Huang、L. Zhang 和 T. Zeng，《用于单图像超分辨率的Transformer》，载于2022年IEEE/CVF计算机视觉与模式识别会议研讨会（CVPRW），2022年，第456–465页。[20] H. Wang、X. Chen、B. Ni、Y. Liu 和 J. Liu，《用于轻量级图像超分辨率的全方位聚合网络》，载于IEEE/CVF计算机视觉与模式识别会议论文集，2023年，第22 378–22 387页。[21] F. Yang、H. Yang、J. Fu、H. Lu 和 B. Guo，《学习用于图像超分辨率的纹理变换器网络》，载于IEEE/CVF计算机视觉与模式识别会议论文集，2020年，第5791–5800页。

[22] H. Chen、Y. Wang、T. Guo、C. Xu、Y. Deng、Z. Liu、S. Ma、C. Xu、C. Xu 和 W. Gao，《预训练图像处理变换器》，载于IEEE/CVF计算机视觉与模式识别会议论文集，2021年，第12 299–12 310页。[23] J. Cao、Y. Li、K. Zhang 和 L. V. Gool，《视频超分辨率变换器》，2023 [ ] 2106 06847。[24] 在线。链接：https://arxiv.org/abs/。J. Liang、J. Cao、G. Sun、K. Zhang、L. Van Gool 和 R. Timofte，《SwinIR：使用Swin Transformer进行图像复原》，载于IEEE/CVF国际计算机视觉会议论文集，2021年，第1833–1844页。[25] Z. Wang、X. Cun、J. Bao、W. Zhou、J. Liu 和 H. Li，《Uformer：通用U形图像复原Transformer》，载于IEEE/CVF计算机视觉与模式识别会议论文集，2022年，第17 683–17 693页。[26] C.-C. Hsu、C.-M. Lee 和 Y.-S. Chou，《DRCT：通过摆脱信息瓶颈实现图像超分辨率》，载于IEEE/CVF计算机视觉与模式识别会议（CVPR）研讨会论文集，2024年6月，第6133–6142页。[27] Q. Huang、A. Jain、G. Stockman 和 A. Smucker，《植物根系结构的自动图像分析》，载于第11届IAPR国际模式识别会议论文集。第二卷。B会议：模式识别方法与系统，1992年，第569–572页。[28] Y. Chen 和 X. Zhou，《基于二维扫描仪的植物根系图像处理与分析》，载于2010年IEEE第五届生物启发计算理论与应用国际会议（BIC-TA）。IEEE，2010年，第1216–1220页。[29] A. French、S. Ubeda-Tom´ás、T. J. Holman、M. J. Bennett 和 T. Pridmore，《利用新型图像分析工具对根系生长进行高通量定量分析》，《植物生理学》，第150卷，第4期，第1784–1795页，2009年。 [在线]。链接：https://academic.oup.com/ plphys/article/ 150 4 1784 6107990 / / / D. Mishra、S. Chemweno、O. [30] Hadar、O. Ben-Tovim、N. Lazarovitch 和 J. E. Ephrath，《用于根毛增强和改善根系性状测量的无监督图像超分辨率》，IEEE 农食电子学汇刊，第2卷，第1期，第81–90页，2024年。[31] L. Lantini、F. Massimi、F. Tosti、A. M. Alani 和 F. Benedetto，《基于深度学习的探地雷达频谱图像树木根系检测方法》，载于2022年第45届电信与信号处理国际会议（TSP），2022年，第391–394页。[32] S. Zhu、Y. Su 和 M. Sato，《基于探地雷达和高精度定位系统的树木根系检测与根系生物量估算》，载于2013年IET国际雷达会议，2013年，第1–5页。[33] W. Luo、Y. H. Lee、M. L. M. Yusof 和 A. C. Yucel，《用于热带地区树木根系有效探测的深度自适应滤波方法》，IEEE仪器与测量汇刊，第72卷，第1–10页，2023年。[34] W. Luo、Y. Hui Lee、T. Hao、M. Lokman Mohd Yusof 和 A. C. Yucel，《用于增强三维树木根系结构重建的自动双极化探地雷达》，IEEE地球科学与遥感汇刊，第62卷，第1–18页，2024年。[35] L. Lantini、I. Giannakis、F. Tosti、D. Mortimer 和 A. M. Alani，《基于反射率的探地雷达信号处理方法：绘制行道树根系分布图》，载于2020年第43届电信与信号处理国际会议（TSP），2020年。[36] S. Zhu、C. Huang、Y. Su 和 M. Lu，《基于探地雷达圆形测量的树木根系检测》，载于第15届国际探地雷达会议论文集，2014年，第135–139页。[37] G. Bodner、M. Alsalem、A. Nakhforoosh、T. Arnold 和 D. Leitner，《用于植物表型分析和生理学研究的RGB与光谱根系成像：实验设置与成像方案》，《可视化实验杂志》，第126卷，第e56251页，2017年。[38] P. K. Jayapal、E. Park、M. A. Faqeerzada、Y.-S. Kim、H. Kim、I. Baek、M. S. Kim、D. Sandanam 和 B.-K. Cho，《利用深度学习分析RGB植物图像以识别韩国人参植株根腐病》，《应用科学》，第12卷，第5期，第2489页，2022年。[39] W. Lu、X. Wang 和 W. Jia，《基于深度学习和先验知识的根毛图像处理》，《农业中的计算机与电子学》，第202卷，第107397页，2022年。[40] S. Zhang、W. Lantz、I. Wilson、J. Spittle、S. Eastin、A. Mandal、J. Arnold 和 T. Bourlai，《一种新的基于样条的图像处理技术及其在植物根系全切片成像中的应用》，载于2024年SoutheastCon会议，2024年，第890–891页。[41] C. Sun、Y. Xu、D. Shen 和 T. Nie，《用于实体匹配中域适应的匹配特征分离网络》，载于2024年ACM网络会议论文集，2024年，第1975–1985页。[42] X. Wu、K. Zhang、Y. Hu、X. He 和 X. Gao，《用于图像超分辨率的多尺度非局部注意力网络》，信号处理，第 218 卷，109362 页，2024。[43] D. P. Kingma 和 J. Ba，《Adam：一种随机优化方法》，2017。[44] K. Zhang、J. Liang、L. Van Gool 和 R. Timofte，《面向深度盲图像超分辨率的实用退化模型设计》，载于 *IEEE*国际计算机视觉会议，2021，第 4791–4800 页。