doi:10.6041/j.issn.1000-1298.2025.10.046

基于 CA－SegNet 网络的灌浆期小麦关键器官分割方法

田世杰[1,2]　闫佳新[1]　秦伟杰[1]　白绍兴[1]　洪旻珂[3,4]　孙章彤[1,3]　胡　瑾[1,3]

（1. 西北农林科技大学信息工程学院，陕西杨凌 712100；2. 作物抗逆与高效生产全国重点实验室，陕西杨凌 712100；
3. 农业农村部农业物联网重点实验室，陕西杨凌 712100；4. 西北农林科技大学机械与电子工程学院，陕西杨凌 712100）

摘要：灌浆期是影响小麦产量与品质的关键时期，其旗叶与麦穗准确识别对表型分析和高产小麦选育具有重要意义。然而，灌浆期小麦点云存在数据量少、器官密集、遮挡严重等问题，为关键器官语义分割带来巨大挑战。为解决这一问题，本研究提出了一种点云数据增强策略和关键器官分割网络（Class-aware segmentation network，CA－SegNet）。首先，基于三维高斯飞溅（3D Gaussian splatting，3DGS）对 80 株小麦进行了三维重建并获得稠密点云数据。其次，为增强样本多样性并保持关键几何结构，提出了动态体素栅格化随机采样（Dynamic voxel farthest point sampling，DVFPS）方法，结合动态体素划分与随机最远点采样策略实现了数据集 10 倍扩增与标准化点云数量（8 192 点）处理。最后，以标准点云数据集为输入，提出了改进的点云分割网络 CA－SegNet，通过类别感知特征提取（Class-aware feature extraction，CAFE）模块和结合分类准确性、空间一致性、特征约束的改进损失函数，有效提高了关键器官分割性能。试验结果表明，CA－SegNet 的平均交并比、精确率、召回率和 F1 值分别达 62.28%、76.41%、72.63% 和 74.47%，均优于 PointNet＋＋、PlantNet 等方法。同时，消融试验也验证了 DVFPS、CAFE 和改进损失函数在提升模型性能方面的有效性。基于分割结果计算的麦穗体积，相关系数 R 和均方根误差分别为 0.85 和 79.16 cm^3，进一步验证了该方法在表型解析方面的有效性。本研究有效提高了小麦点云中关键器官分割精度，为智慧农业中小麦表型检测与精准育种提供了可靠技术手段。

关键词：灌浆期小麦；三维点云；语义分割；点云数据增强；改进损失函数；植物表型

中图分类号：S126　　**文献标识码**：A　　**文章编号**：1000-1298(2025)10-0520-10　　**OSID**：

Key Organ Segmentation Method for Wheat at Grain Filling Stage Based on CA－SegNet Network

TIAN Shijie[1,2]　YAN Jiaxin[1]　QIN Weijie[1]　BAI Shaoxing[1]　HONG Minke[3,4]　SUN Zhangtong[1,3]　HU Jin[1,3]

(1. *College of Information Engineering*, *Northwest A&F University*, *Yangling*, *Shaanxi* 712100, *China*
2. *State Key Laboratory for Crop Stress Resistance and High-Efficiency Production*, *Yangling*, *Shaanxi* 712100, *China*
3. *Key Laboratory of Agricultural Internet of Things*, *Ministry of Agriculture and Rural Affairs*, *Yangling*, *Shaanxi* 712100, *China*
4. *College of Mechanical and Electronic Engineering*, *Northwest A&F University*, *Yangling*, *Shaanxi* 712100, *China*)

Abstract: The grain filling stage is a critical period affecting wheat yield and quality, where accurate identification of flag leaves and spikes holds significant importance for phenotypic analysis and high-yield wheat selection. However, wheat point clouds during this stage face challenges such as sparse data, dense organ distribution, and severe occlusions, posing substantial difficulties for semantic segmentation of key organs. To address these issues, a point cloud data augmentation strategy and a key organ segmentation network, termed class-aware segmentation network (CA－SegNet) were proposed. Firstly, dense point cloud data were acquired through 3D reconstruction of 80 wheat plants using the 3D Gaussian splatting (3DGS) algorithm. Secondly, to enhance sample diversity while preserving critical geometric structures, a dynamic voxel farthest point sampling (DVFPS) method was developed, which integrated dynamic voxel partitioning with farthest point sampling to achieve a 10-fold dataset expansion and

收稿日期：2025－04－18　修回日期：2025－06－05

基金项目：陕西省自然科学基础研究计划项目（2025JC－YBQN－224）、杨凌示范区科技计划项目（2024NY－06）、陕西省重点研发计划项目（2024NC2－GJHX－36）、旱区农业陕西实验室建设开放课题（2024ZY－JCYJ－02－16）和西北农林科技大学博士科研启动费项目（2452024119）

作者简介：田世杰（1994—），男，副教授，博士，主要从事多模态表型信息感知与智能装备等研究，E-mail：sjtian@nwafu.edu.cn

通信作者：胡瑾（1980—），女，教授，博士，主要从事农业人工智能与智能决策等研究，E-mail：hujin007@nwsuaf.edu.cn

standardized point cloud size (8 192 points). Finally, the improved CA－SegNet was proposed, which introduced a class-aware feature extraction (CAFE) module in the encoder stage. This module incorporated dual-stage class attention (DCA) to fuse spatial structural information, enabling prioritized focus on semantically significant central points. Additionally, an enhanced loss function combining classification accuracy, spatial consistency, and feature constraints was designed to further optimize segmentation performance. Experimental results demonstrated that CA－SegNet achieved superior performance with mIoU, precision, recall, and F1-score of 62.28%, 76.41%, 72.63%, and 74.47%, respectively, outperforming methods such as PointNet++ and PlantNet. Ablation studies confirmed the effectiveness of DVFPS, CAFE, and the improved loss function in enhancing model performance. The spike volume calculated based on the segmentation results achieved an R value of 0.85 and an RMSE of 79.16 cm^3, further validating the effectiveness of the method in phenotypic analysis. The research significantly improved the segmentation accuracy of key organs in wheat point clouds, providing a reliable technical approach for wheat phenotypic detection and precision breeding in smart agriculture.

Key words: wheat at grain filling stage; 3D point cloud; semantic segmentation; point cloud data augmentation; enhanced loss function; plant phenotype

0　引言

小麦作为世界三大粮食作物之一，为全球约35%的人口提供口粮[1]。提高小麦产量对于应对全球人口增长带来的粮食安全挑战具有重要战略意义。研究表明，在小麦生长过程中，旗叶与麦穗是决定产量的关键部位[2-5]。旗叶是小麦光合作用的主要贡献者，而麦穗直接影响籽粒形成与品质[6-8]。因此，对小麦旗叶和麦穗等关键器官进行语义分割和精准识别，可为小麦表型分析和高产优质小麦品种培育提供关键技术支持。

传统的植物表型参数测量方法依赖人工测量与近似计算，存在误差大、效率低和破坏性等问题，难以满足大规模、高通量的表型测量需求。近年来，计算机视觉和深度学习的融合，在二维图像表型分析中得到了广泛应用[9-10]。然而二维图像缺乏深度信息，难以处理复杂遮挡问题。相比之下，三维点云不仅包含更丰富的纹理信息，还具备最关键的深度信息，有效突破了二维图像的瓶颈。KOLODIAZHNYI 等[11]通过将不规则结构点云转换为规则的体素网格，并基于三维卷积网络实现了家具语义分割。XIANG 等[12]通过将点云简化为拓扑骨架，揭示了高粱茎叶分蘖的层级结构。QI 等[13]提出 PointNet 深度学习框架，通过置换不变性处理机制实现了无序点云分割的端到端框架。LI 等[14-15]在 PSegNet 和 PlantNet 网络中引入空间与通道注意力机制，为点云分割提供了新思路。尽管当前作物的点云分割研究已取得一定进展[16]，但对于小麦复杂点云数据，尤其是灌浆期小麦点云，在语义分割时仍面临困难：①传统点云下采样方法在复杂植株场景中易导致关键器官特征丢失，影响后续分割精度。②现有分割网络在空间特征与语义信息联合建模方面仍有不足，难以应对小麦冠层中密集叶片与麦穗的相互遮挡问题。

针对灌浆期小麦复杂点云数据关键器官语义分割问题，本研究提出一种数据增强策略和改进的分割网络框架 CA－SegNet，提出 DVFPS 数据增强方法，在扩展数据多样性的同时保持关键几何特征；在编码阶段设计 CAFE 模块，通过融合 DCA 与加权最远点采样策略，改善采样时忽略重要点的问题；提出一种结合点云空间结构与语义信息的改进损失函数，以有效提升小麦复杂点云数据语义分割精度。

1　小麦复杂点云数据

1.1　数据采集

选用西农 877 和小偃 22 小麦品种作为试验材料，各设置 40 盆培育样本，共计 80 盆。所有样本均栽培于西北农林科技大学作物抗逆与高效生产全国重点实验室科研温室(34°07′39″N，107°59′50″E，海拔 648 m)。针对小麦复杂表型特征及植株间相互遮挡问题，建立多视角图像采集系统：采用环绕拍摄方式，分别从植株冠层上部、穗部中部和基部下部 3 个观测视角获取视频数据，通过帧率 2 f/s 采样提取序列图像，构建覆盖植株 360°立体空间的多视角图像数据集。

基于该数据集，采用 3DGS[17]算法实现小麦植株的三维重建(图 1)，该方法通过可微分渲染技术有效解决植物器官自遮挡导致的点云缺失问题。重建过程中，关键参数设置如下：迭代次数为 30 000，球谐函数阶数为 3，不透明度学习率为 0.025，其余参数保持默认值。

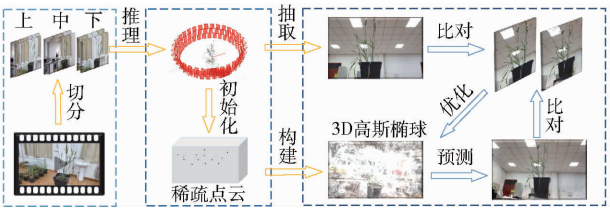


图 1 灌浆期小麦三维重建流程

Fig. 1 3D reconstruction process of wheat during grain filling period

1.2 数据标注

基于 CloudCompare 工具对小麦点云数据进行精确标注。在语义标注前,由于原始小麦点云数据是通过多视角视频重建所生成,会伴随除小麦主体以外的复杂背景干扰。因此,需进行手动分割处理:提取小麦主体区域并剔除离群点。在语义标注阶段,将小麦点云划分为 3 类:麦穗(类别标签:0)、旗叶(类别标签:1)以及其他类别(类别标签:2)。

为保证数据集质量,引入峰值信噪比(Peak signal-to-noise ratio, PSNR)对三维重建点云质量进行评估。通常认为当 PSNR 高于 30 dB 时,重建质量较好[18]。而当 PSNR 低于 20 dB 时,点云中小麦目标结构模糊,难以从场景中准确分辨出小麦形态。因此,对所有样本点云重建结果进行筛选,剔除其中 PSNR 低于 20 dB 的 5 个低质量样本,最终保留 75 个高质量样本(平均 PSNR 为(35.23 ± 1.71) dB)用于后续试验。

1.3 数据增强

在建立深度学习网络前,需考虑数据集的多样性及网络输入要求。一方面,试验初期所建立的 75 个小麦点云数据集样本量少,导致模型在训练过程中易出现过拟合现象。另一方面,点云分割网络通常需对输入数据规模进行限制,以减小网络参数数量,提高计算效率。当前三维点云在作物表型研究中,通常将点云下采样至 4 096 点作为网络输入[14]。虽然更少的点数会提升训练效率,但本研究使用的灌浆期小麦点云数据集结构复杂、信息密集,若点数过少,会导致关键表型特征丢失。为兼顾特征完整性与模型训练需求,将点云下采样至 8 192 点,并结合数据增强策略,以提升数据多样性和模型泛化能力。

目前已有多种点云数据增强方法,例如基于体素化的采样(Voxel-based sampling, VBS)[19]和最远点采样(Farthest point sampling, FPS)[20]。其中,VBS 具有很高的处理速度并且可以尽可能多地保留点云特征,但该算法存在两个缺点:①体素化尺寸需根据点云密度进行调整,这导致不同点云数据采样后点数并不确定,无法直接用于模型训练。②VBS 所选取的采样点是体素的重心点,这可能导致局部密度恒定,生成密度均匀的规则点云,数据不具有多样性,不利于模型训练。FPS 可以保持采样后的点云局部密度并得到固定点数的采样点,但该方法仍有缺点:①在稀疏区域可能会忽视关键特征,在密集区域可能会过度采样冗余点。②采样结果对初始点敏感,难以直接使用 FPS 生成多样化的训练样本。

在比较 VBS 和 FPS 优缺点后,提出一种新型的点云数据增强方法——动态体素栅格化随机采样(DVFPS),结合 VBS 与 FPS 优势,在保留关键几何特征的同时实现数据规模可控性与多样性扩增。具体设计如下。

1.3.1 动态体素化下采样

引入八叉树层级自适应机制(图 2),该机制通过点云最大包围盒尺寸以及八叉树层级可动态计算出每个点云体素尺寸,计算式为

$$l_{Cell_size} = \frac{L}{2^l} \tag{1}$$

式中 l_{Cell_size}——体素边长

L——点云最大包围盒边长

l——八叉树层级

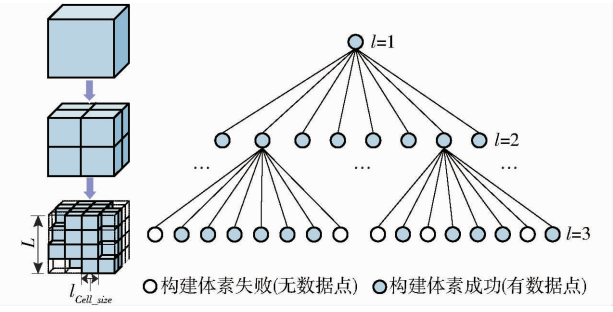


图 2 点云八叉树构建体素过程

Fig. 2 Voxel construction process of point cloud octree

式(1)通过空间递归划分[21],确保体素分辨率与点云局部密度自适配,显著减少采样后点数波动性。具体实现流程为:①体素网格构建:设置八叉树层级为 8,计算出 l_{Cell_size},将原始点云划分为均匀立方体体素网格,每个体素保留其内部所有点的坐标与语义信息。②语义感知重心采样:对每个体素内点云坐标取均值生成重心点,语义标签由多数投票原则确定[22],避免传统重心法导致的语义模糊问题。此阶段将点云规模压缩至约 50 000 点,在保留植株旗叶与穗部关键结构的同时,为后续精细化采样提供高密度特征基础。

1.3.2 随机最远点采样

提出以下改进策略:①多候选点随机选择:在每次迭代中,筛选距离当前采样点集最远的 K 个候选点($K = 10$),并从中随机选取下一采样点,缓解初始

点依赖性。②局部密度均衡约束：通过计算候选点邻域密度权重，优先选择稀疏区域特征点，避免密集区域冗余采样[20]。③多样性扩增机制：对每份动态体素化下采样结果执行 10 次独立的随机最远点采样，保证每次选取的初始点随机且不重复，以此确保数据的多样性。最终生成 10 组 8 192 点的点云数据，模拟真实场景中点云的空间分布不确定性[23]。

最终通过 DVFPS 方法（图 3），原始 75 盆小麦点云数据被扩增至 750 组标准化点云（每组 8 192 点），有效缓解数据稀缺性与过拟合风险。

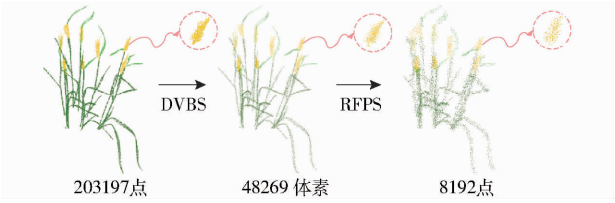


图 3　小麦点云数据下采样流程

Fig. 3　Sampling process of wheat point cloud data

1.4　数据增强评价指标

在三维点云处理中，数据增强操作需在下采样处理过程的同时保留关键几何与语义信息。本节构建多维度几何与语义评估体系定量分析采样质量。

1.4.1　几何结构保留

倒角距离（Chamfer distance，CD）[24]是衡量 2 个点云间几何相似性的重要指标，尤其在下采样任务中得到广泛应用，该指标通过双向最近邻搜索量化全局几何偏差。豪斯多夫距离（Hausdorff distance，HD）[25]用于衡量 2 个点云数据间的局部最大几何偏差，尤其关注点云间最不匹配的区域（如小麦旗叶边缘）。其核心思想是计算 2 个点云数据间“最远邻近点”的最大距离，反映局部结构的保真度。CD 通过双向最近邻距离平均值，反映点云整体的几何一致性。其归一化特性使不同规模点云的评估结果可比。HD 则捕捉点云间的最远邻近距离，对局部异常值高度敏感，能体现不同采样点数在细粒度结构上的差异。二者结合可评估点云下采样对小麦器官曲率、拓扑连接的保留能力。

1.4.2　语义分布一致性验证

L1 分布距离又称曼哈顿距离，用于衡量 2 个概率分布间的绝对差异，适用于评估下采样前后点云语义标签分布的一致性。KL 散度（Kullback-leibler divergence，KLD）[26]是衡量 2 个概率分布差异的重要指标，针对本研究数据可用于量化下采样结果的信息损失。

L1 分布距离直接度量语义标签分布的绝对差异，KLD 则从信息熵角度反映分布间的不对称差异[27]。该指标体系对麦穗和旗叶交叠区等细粒度语义边界的采样敏感性具有重要诊断价值。

2　关键器官分割方法

2.1　CA－SegNet 网络

基于 PointNet＋＋[20]传统的编码-解码结构，提出一种改进的点云分割网络 CA－SegNet（图 4），优化小麦点云数据的特征提取过程。

在编码阶段，基于多层级的采样-分组-聚合模块（Set abstraction，SA）[20]提出了改进类别感知的特征提取模块，以完成点云的层次化局部特征提取。其中，在下采样过程中，按照点的重要性得分选择中心点；在分组过程中，以多尺度半径查询策略[20]，增强对小麦点云非均匀密度分布的适应性；在特征提取过程中，通过 MLP 和最大池化操作获得局部特征。在解码阶段，通过特征传播（Feature propagation，FP）模块[20]结合跳跃连接[28]结构，逐步恢复点云分辨率。在特征传播过程中，采用 KNN 的加权插值策略，将当前层特征向上传递至上一层，以更精细地恢复局部几何细节。同时，将插值特征与上一层原始特征进行拼接，通过 MLP 完成特征融

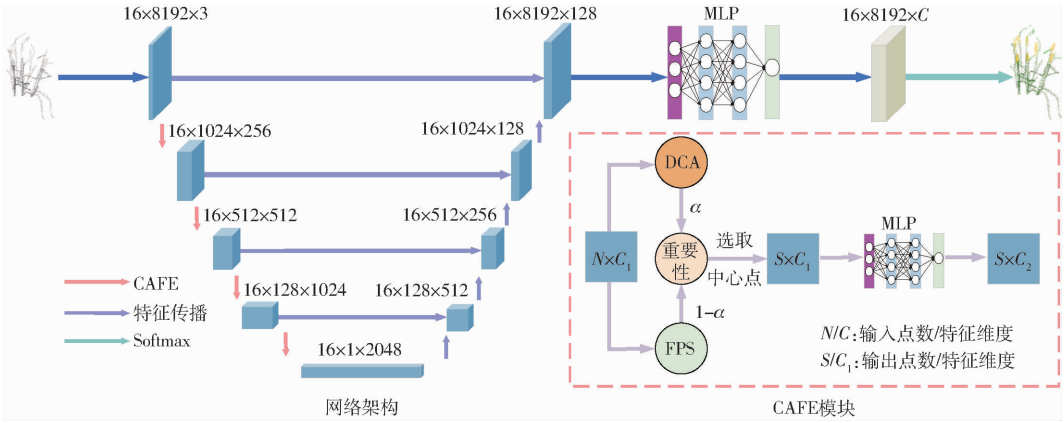


图 4　CA－SegNet 网络架构

Fig. 4　CA－SegNet network architecture

合。最终输出每个点的分类概率,实现点级语义分割。

本研究的核心创新在于设计了 CAFE 模块,融合双阶段类别注意力机制与加权最远点采样策略,在保证空间均匀性的同时,增强了对语义关键区域的关注,从而提高了小麦点云分割效果。

2.2 类别感知局部特征提取模块

在传统特征提取模块中,由于基于 FPS 选择中心点的方法仅依赖点云空间分布[20],因此很容易忽略一些具有重要语义的特征点。ZHU 等[29]提出了一种用于多标签分类任务的类别特定残差注意力(Class-specific residual attention, CSRA)模块,其通过类别特定的注意力机制增强特征表达能力。受此启发,本研究提出 CAFE 模块,扩展了原始工作,并实现了从二维图像到三维点云的适应性改进。

2.2.1 双阶段类别注意力机制

双阶段类别注意力机制融合了多头自注意力[30]和类别特定注意力(图 5)。

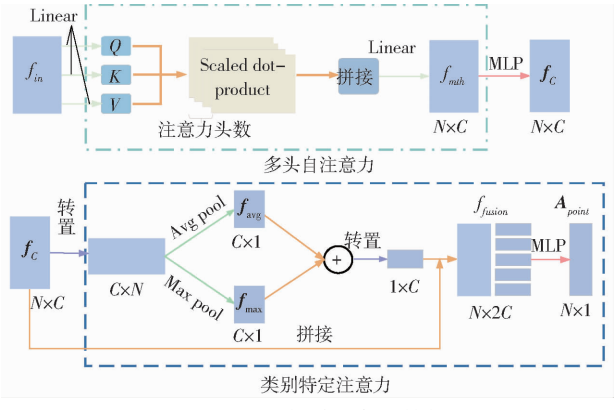


图 5 双阶段类别注意力结构图

Fig. 5 Structure of two-stage category attention module

根据点云数据中的不同类别特征,为每个点生成类别注意力得分。多头注意力能捕捉到点云数据中的多样化、复杂的局部和全局信息,从而使模型更全面地理解点云中的结构信息。而类别特定注意力能充分利用每个类别单独的空间注意力,为点云中的每个点生成不同的类别注意力得分,用以选择具有更高语义价值的中心点。

首先,对输入的点云特征 f_{in} 通过多头自注意力层建模全局点间关系,得到特征 f_{mth}。为适应类别特定注意力的输入,将特征 f_{mth} 通过 MLP 映射至类别空间 $\mathbf{R}^{N\times C}$,得到特征 $\boldsymbol{f}_C$。之后,采用双池化融合策略,计算类别特定注意力。其中,平均池化提取了类别的整体信息,而最大池化关注于该类别最具代表性的点,分别得到特征向量 $\boldsymbol{f}_{\text{avg}}$ 和 $\boldsymbol{f}_{\max}$。将平均池化和最大池化的结果相加即可得到全局类别注意力 $\boldsymbol{A}_C$,计算式为

$$\boldsymbol{A}_C = \boldsymbol{f}_{\text{avg}} + \boldsymbol{f}_{\max} \tag{2}$$

其中

$$\boldsymbol{f}_{\text{avg}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{f}_{C,i}^{\mathrm{T}} \tag{3}$$

$$\boldsymbol{f}_{\max} = \max(\boldsymbol{f}_{C,0}^{\mathrm{T}}, \boldsymbol{f}_{C,1}^{\mathrm{T}}, \cdots, \boldsymbol{f}_{C,N}^{\mathrm{T}}) \tag{4}$$

式中 C——类别数

由于最大池化获取了每个类别的所有空间位置中的最大值,因此可以看作一种类别特定的注意力机制[29]。通过拼接 $\boldsymbol{A}_C$ 和 $\boldsymbol{f}_C$ 获取每个点的类别特征 $f_{fusion,i}$,之后通过 MLP 得到每个点最终的类别注意力 $\boldsymbol{A}_{point}$。最后,对 $\boldsymbol{A}_{point}$ 归一化得到每个点的注意力得分 $\boldsymbol{A}_{norm}$。

2.2.2 加权最远点采样策略

基于传统的中心点选取方法,提出了一种注意力引导的最远点采样方法。该方法通过融合类别注意力得分与几何距离得分以衡量点的重要性,实现语义与空间协同优化的采样策略。算法流程如下:

输入:点云坐标 X;类别注意力得分 A_norm;
采样点数 S;平衡系数 α

输出:采样点集合

1: p_0 = center // 随机选择初始中心点
2: D_min = $+\infty$ // 初始化全局最小距离矩阵
3: Samples = [ ] //采样点集合
4: for k in range(1, S + 1) do:
5: $D_current$ = $p_(k-1)$到所有点的欧氏距离 // 当前中心点到每个点的欧氏距离
6: D_min = Min(D_min, $D_current$) // 更新全局最小距离
7: D_norm = Norm(D_min) // 对全局最小距离进行归一化
8: I = αA_norm + (1 − α)D_norm // 计算重要性得分
9: p_k = Max(I) // 选择新的采样点
10: Samples. append(p_k) // 更新采样点集合
11: end for
12: return Samples //得到采样点集合

小麦点云中每个点重要性得分定义为

$$I_i = \alpha A_{norm,i} + (1-\alpha) D_{norm,i} \tag{5}$$

式中 $D_{norm,i}$——归一化后点 i 到当前已选择点集的最短欧氏距离

通过双阶段类别注意力模块机制与加权最远点采样策略,该模块将语义特征与空间分布信息有效融合,从而优化了局部特征提取过程。

2.3 改进损失函数设计

为进一步提升小麦点云语义分割效果,提出了

一种复合损失函数 L_{CIR}，综合考虑分类准确性、空间一致性与特征约束。该损失函数由 3 部分组成：交叉熵（Cross-entropy）损失、完整交并比（Complete intersection over union，CIoU）损失[31]以及特征正则化（Feature regularization）损失。最终总损失函数表达为

$$L_{CIR}=\beta L_{CE}+\gamma L_{\mathrm{CIoU}}+\delta L_{reg} \tag{6}$$

其中，β、γ 和 δ 分别为各项损失权重，为确保各项损失在训练过程中对整体优化过程发挥适当作用，固定 $\beta=1.0$，$\gamma=1.0$，$\delta=0.01$。该设置综合考虑各损失项在训练中的量级差异，避免某一项因数值过大或过小而主导或弱化模型学习方向。本损失函数的设计不仅提升了分割精度，同时增强了点云特征提取稳定性，并在 3.3 节证明了其有效性。

交叉熵损失用于衡量模型预测类别与真实类别之间的差异。该损失函数通过最小化模型预测和真实类别之间的差异，促进网络学习到更准确的类别分布。传统的交并比（Intersection over union，IoU）主要用于检测任务，但在点云分割任务中，IoU 单纯的重叠度度量可能不足以反映分割质量。为此，引入 CIoU[31]，它不仅计算 IoU，还考虑类别分布中心偏移和形状相似性。

通过引入类别分布的中心偏移和形状相似性，CIoU 损失提升了点云分割的稳定性和准确性，特别是在处理小型物体和密集区域时，能有效避免常见的分割不一致问题。

特征正则化损失用于约束模型的特征空间，防止过拟合并增强特征提取的鲁棒性。对全局特征向量 $\boldsymbol{f}\in\mathbf{R}^{d}$ 施加 L2 正则化。通过引入特征正则化，可使模型在训练过程中避免过度拟合，并提升模型泛化能力。

2.4　模型训练与测试

基于 PyTorch 框架构建模型，并在 NVIDIA RTX 4090D GPU（24 GB 显存）及 Intel（R） Xeon（R） Platinum 8481C（16vCPU）平台上进行训练，以确保计算的高效性与稳定性。使用自建小麦点云数据集作为试验数据集。为确保评估客观性与训练样本独立性，从 75 个原始样本中随机选取 10 个作为测试集来源，其对应的 100 个增强样本构成测试集，其余 650 个增强样本用于模型训练。模型训练总共进行 150 轮（Epochs），批量大小（Batch size）经多次试验后最终设定为 16，并采用 Adam 优化器进行参数更新，以提高训练的收敛速度和稳定性。

为评估模型语义分割性能，采用平均交并比（mIoU）、精确率（Precision）、召回率（Recall）和 F1 值作为主要评价指标。

3　试验结果与分析

3.1　数据增强方法评估

为验证所提出的 DVFPS 方法的有效性，基于 1.4 节所建立的评价体系，对 VBS、FPS、三维边缘保持采样（3D edge-preserving sampling，3DEPS）[15]和 DVFPS 的数据增强结果进行对比（图 6）。在对比试验中，VBS 先采用固定体素尺寸（0.01）进行降采样，之后使用随机算法定点采样 8 192 点，FPS 以及 DVFPS 则直接按固定 8 192 点采样；3DEPS 依照推荐比例，边缘点占总点数的 30%，内部点占总点数的 70%。所有试验在相同环境下完成。

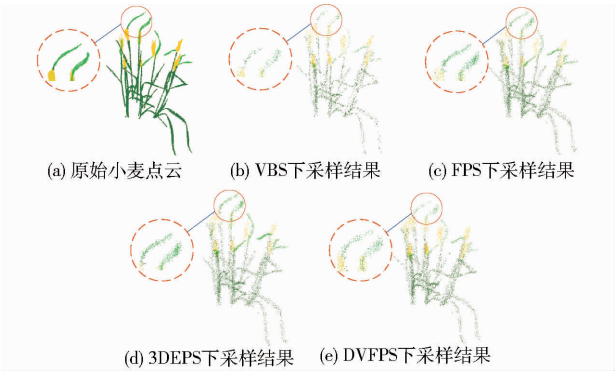


图 6　下采样结果对比

Fig. 6　Comparison of down sampling results

4 种采样方法对比结果如表 1 所示。试验结果表明，DVFPS 在几何保真度指标上显著优于其余算法。DVFPS、FPS 和 3DEPS 的 CD 指标均达 0.000 1，较 VBS（0.000 3）降低 66.67%，表明三者在全局几何一致性上更优。HD 指标进一步揭示局部特征差异：DVFPS 的 HD 为 0.022 5，较 FPS（0.031 4）降低 28.34%，较 VBS（0.137 3）降低 83.61%，较 3DEPS（0.043 7）降低 48.51%。

表 1　不同采样方法对比结果

Tab. 1　Comparison results of different sampling methods

采样方法	CD	HD	L1	KLD
VBS	0.000 3	0.137 3	0.171 9	0.036 9
FPS	0.000 1	0.031 4	0.185 6	0.032 0
3DEPS	0.000 1	0.043 7	0.162 5	0.028 2
DVFPS	0.000 1	0.022 5	0.158 0	0.026 1

语义标签的分布一致性直接影响模型对器官边界的识别精度，试验结果表明：DVFPS 在 L1 分布距离与 KLD 上均表现最优。其 L1 为 0.158 0，较 VBS（0.171 9）、FPS（0.185 6）和 3DEPS（0.162 5）分别降低 8.09%、14.87% 和 2.77%；KLD 为 0.026 1，较 VBS（0.036 9）、FPS（0.032 0）和 3DEPS（0.028 2）分别降低 29.27%、18.44% 和 7.45%。上述结果显示本研究所提出的 DVFPS 在几何保留、语义一致性方

面均展现显著优势,适用于后续模型训练。

3.2 模型分割结果

为验证模型分割效果,在自建数据集上评估了麦穗、旗叶及其他器官的分割性能,试验结果如表 2 所示。本研究方法在多个评估指标上均取得了良好结果,其中麦穗区域平均交并比达 66.16%,表明所提出方法能有效识别小麦穗部结构。同时,其他器官平均交并比也较高,说明模型具备良好的整体分割能力。

表 2 模型评估结果

Tab. 2 Model evaluation results %

类别	平均交并比	精确率	召回率	F1 值
麦穗	66.16	82.99	76.54	79.63
旗叶	34.55	56.70	46.91	51.34
其他器官	86.13	89.86	95.41	92.55
平均	62.28	76.41	72.63	74.47

尽管本研究方法在整体分割效果上取得了较好的结果,但由于旗叶与其他叶片仅存在位置上的差异,且小麦点云数据中旗叶、麦穗和其他器官存在一定程度的相互遮挡和交叉,这导致旗叶区域的分割效果相对较低,平均交并比仅为 34.55%。而其他部分(非旗叶和麦穗区域)由于点云数量占比最高,所以该部分的平均交并比会更高。麦穗的平均交并比则介于两者之间。

3.3 消融试验

为验证网络中所使用的 DVFPS、CAFE 和 L_{CIR} 改进损失函数设计对模型的贡献,试验中分别移除不同模块进行消融试验,结果如表 3 所示。由表 3 可以发现,移除 DVFPS、CAFE 和 L_{CIR} 改进损失函数都会使网络分割效果变差,在平均交并比、精确率、召回率和 F1 值上分别平均下降 2.54%、1.76%、1.03%、1.39%。由此可见,本研究提出的 DVFPS、CAFE 及优化 Loss 设计在提升点云分割效果方面均具有重要作用,能够有效提高小麦点云分割精度和模型泛化能力。

表 3 消融试验评估结果

Tab. 3 Assessment results of ablation experiment %

DVFPS		CAFE	L_{CIR}	平均交并比	精确率	召回率	F1 值
动态体素化	随机最远点采样						
×	×	×	×	59.44	73.53	69.83	71.63
√	×	√	√	60.60	74.77	72.19	73.46
×	√	√	√	61.65	75.94	72.33	74.09
√	√	×	×	60.23	74.25	71.38	73.32
√	√	√	×	60.78	74.94	72.02	73.45
√	√	×	√	60.89	75.74	71.71	73.67
×	×	√	√	60.32	74.51	71.92	73.19
√	√	√	√	62.28	76.41	72.63	74.47

注:√表示采用此设计,×表示未采用此设计。

3.4 与其他方法对比

为验证本研究提出方法的有效性,将其与现有的点云分割方法进行了对比,包括 PointNet[13]、PointNet++[20]、PlantNet[15]和 PSegNet[14]等。这些网络均使用与本研究相同的输入点云数量(8 192),并采用各自的推荐参数进行试验。试验结果如表 4 所示。

表 4 模型效果对比

Tab. 4 Comparison of model effects %

网络	平均交并比	精确率	召回率	F1 值
PointNet	49.61	61.93	59.22	60.54
PointNet++	59.44	73.53	69.83	71.63
PlantNet	53.43	66.90	64.24	64.54
PSegNet	57.03	71.32	67.22	67.82
Point cloud Transformer[32]	62.16	77.05	72.03	74.45
PAConv[33]	59.89	74.48	70.31	72.34
GDANet[34]	61.97	74.12	74.48	74.30
CA-SegNet(本研究)	62.28	76.41	72.63	74.47

由表 4 可以看出,针对灌浆期小麦点云数据集,本研究方法在平均交并比、精确率、召回率和 F1 值 4 个评估指标上均达到最优或接近最优,显示出其良好的综合性能。为了更直观地展示不同分割网络性能差异,图 7a、7b 分别为使用不同方法在常规情况下和叶穗交叠相互遮挡情况下的分割结果,从图

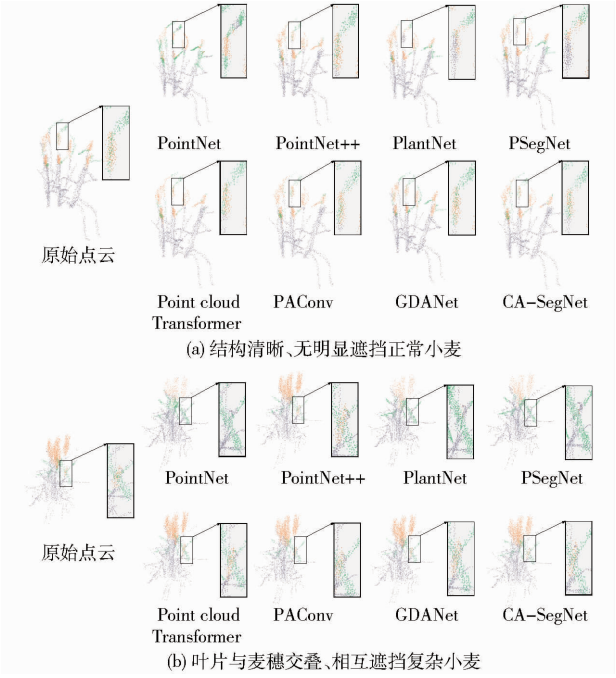


图 7 小麦点云数据语义分割结果可视化

Fig. 7 Visualization of semantic segmentation results of wheat point cloud data

中可以清晰地看到 CA－SegNet 在精度和细节恢复方面的优势。这说明了 CA－SegNet 在处理复杂的器官遮挡和密集结构时，能更好地捕捉到关键特征。

4　讨论

4.1　加权最远点采样平衡系数设置

平衡系数在加权最远点采样中起着至关重要的作用，它决定了类别注意力和几何距离在点选择中的相对影响。为优化局部特征提取中中心点的选择，注意力加权的最远点采样策略综合考虑类别注意力权重 A_{norm} 和几何距离权重 D_{norm}，通过加权公式(5)计算每个点的重要性。为了确定最优平衡系数，以 0.05 为步长，在不同 α 值下进行试验(图 8)。试验结果表明，当 $\alpha=0.50$ 时，模型综合效果达到最佳，平均交并比最高，分割性能最优。

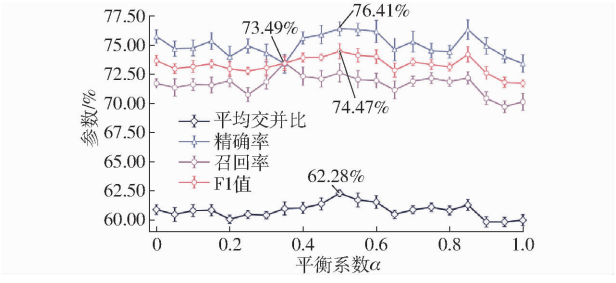


图 8　不同平衡系数下模型精度变化曲线

Fig. 8　Effect of different balance coefficients on model accuracy

4.2　数据扩增倍数设置

数据扩增倍数对提高模型训练效果和泛化能力至关重要。为探究更适宜模型训练的数据扩增倍数，基于 DVFPS 方法生成了多种扩增倍数(5、10、15、20 倍)的训练集，并对比其分割性能。试验结果表明(表 5)，10 倍数据扩增倍数在模型性能与计算成本间取得了最佳平衡：当扩增倍数从 5 倍提升至 10 倍时，平均交并比显著提高 3.23%，其他指标(精确率、召回率、F1 值)均同步上升。这表明 10 倍扩增增加了原始数据的多样性，通过 DVFPS 生成的样本充分覆盖了小麦点云的空间分布特征[23]，增强了模型对旗叶、麦穗等关键结构的辨识能力。然而，当扩增至20倍时，平均交并比相较于10倍下降 0.35%，表明过高的扩增倍数可能引入冗余噪声[20]，尤其在小麦点云遮挡区域(如旗叶与穗部交叉部分)表现更为明显。

表 5　数据扩增倍数效果对比

Tab. 5　Comparison of data amplification factor effects

扩增倍数	数据总量	平均交并比/%	精确率/%	召回率/%	F1 值/%
5	375	60.33	75.54	71.97	73.71
10	750	62.28	76.41	72.63	74.47
15	1 125	62.34	76.52	72.75	74.59
20	1 500	62.06	76.26	72.44	74.30

从训练效率角度分析，扩增倍数与训练耗时呈非线性增长关系。在 NVIDIA RTX 4090D GPU 上，10 倍数据(750 样本)完成 150 轮训练耗时约 1 h，而 15 倍数据(1 125 样本)需约 3 h。因此选择 10 倍扩增作为训练方案，最终，将 75 盆原始小麦点云 10 倍扩增至 750 盆。

4.3　麦穗体积计算

为验证所提出方法在小麦表型参数提取中的实际应用价值，结合凸包体积计算方法，提取每盆小麦的麦穗总体积。具体流程如下：①根据人工测量的小麦株高，对点云数据进行缩放处理，使其与实际高度一致。②采用基于密度的聚类算法(Density-based spatial clustering of applications with noise，DBSCAN)[35]对麦穗点云进行聚类。经过多次测试，设定邻域半径 ε 为 0.05 倍的缩放比例，最小邻域样本点数(min-Points，minPts)为 5 以获得良好的聚类效果。③计算每个聚类子结构凸包体积，累加得到单盆小麦麦穗总体积。

利用测试集所选取的 100 个小麦样本标注数据和 CA－SegNet 网络分割结果，基于上述方法提取每盆小麦麦穗总体积，作为对照值和计算值。通过回归分析(图 9)得到相关系数 R 为 0.85，均方根误差(Root mean qquared error，RMSE)为 79.16 cm^3，参考计算叶倾角中相关系数 R 为 0.84 时认为具有良好精度[36]，本研究在小麦复杂结构条件下对照值与计算值之间的误差在可接受范围内，验证了 CA－SegNet 在小麦表型分析中的有效性。

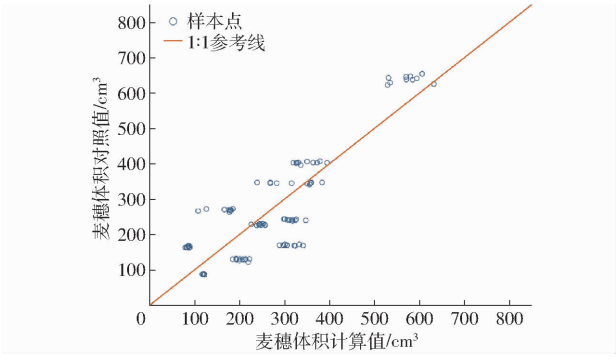


图 9　麦穗体积对照值与计算值回归分析

Fig. 9　Regression analysis chart of predicted and actual values of wheat ear volume

5　结论

(1)针对灌浆期小麦旗叶与麦穗点云语义分割中存在的器官密集遮挡以及数据量较少的问题，提出了

一种数据增强策略和改进的分割模型框架 CA - SegNet。提出的 DVFPS 方法增强了数据多样性,有效降低了模型过拟合风险;改进的 CAFE 模块通过优化特征提取中的中心点选择,提升了模型在复杂场景下的表现;设计的损失函数 L_{CIR} 优化了模型分割精度与鲁棒性。消融试验结果表明,移除 DVFPS、CAFE 和 L_{CIR} 均导致模型性能下降,其中平均交并比、精确率、召回率和 F1 值平均下降 2.54%、1.76%、1.03%、1.39%,验证了各部分的有效性。

(2)试验结果表明,CA - SegNet 在自建小麦点云数据集关键器官分割中有效提升性能。其中,平均交并比为 62.28%,精确率为 76.41%,召回率为 72.63%,F1 值为 74.47%,综合优于 PointNet、PointNet + +、PlantNet 和 PSegNet 等网络。

参 考 文 献

[1] 吴群. 小麦 Unigene 遗传图谱构建和产量性状 QTL 分析[D]. 泰安:山东农业大学,2021.
WU Qun. Construction of Unigene genetic map and QTL mapping for yield traits in wheat[D]. Taian: Shandong Agricultural University, 2021. (in Chinese)

[2] HAMPTON J G, MCCLOY B L, MCMILLAN D R. Ear populations and wheat production[J]. New Zealand Journal of Experimental Agriculture, 1981, 9(2): 185 - 189.

[3] SIDDIQUE K H M, WHAN B R. Ear: stem ratios in breeding populations of wheat: significance for yield improvement[J]. Euphytica, 1993, 73: 241 - 254.

[4] RAHMAN M A, HAQUE M E, SIKDAR B, et al. Correlation analysis of flag leaf with yield in several rice cultivars[J]. Journal of Life and Earth Science, 2013, 8: 49 - 54.

[5] RACZ I, HIRIŞCĂU D, BERINDEAN I, et al. The influence of flag leaf removal and its characteristics on main yield components and yield quality indices on wheat[J]. Agronomy, 2022, 12(10): 2545.

[6] 杨斌,乔玲,赵佳佳,等. 小麦旗叶叶绿素含量的 QTL 定位及验证[J]. 作物学报,2023,49(3):744 - 754.
YANG Bin, QIAO Ling, ZHAO Jiajia, et al. QTL mapping and validation of chlorophyll content in wheat flag leaves[J]. Acta Cropae Sinica, 2023,49(3): 744 - 754. (in Chinese)

[7] 金立桥. 冬小麦麦穗光合生理特性及其对籽粒灌浆的影响[D]. 泰安:山东农业大学,2017.
JIN Liqiao. Photosynthetic physiological characteristics of winter wheat spike and its effect on grain filling[D]. Taian: Shandong Agricultural University, 2017. (in Chinese)

[8] GREGERSEN P L, FOYER C H, KRUPINSKA K. Photosynthesis and leaf senescence as determinants of plant productivity[M]// Biotechnological approaches to barley improvement. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014: 113 - 138.

[9] 董荣胜, 马雨琪, 刘意, 等. 加强类别关系的农作物遥感图像语义分割[J]. 中国图象图形学报, 2022, 27(11): 3382 - 3394.
DONG Rongsheng, MA Yuqi, LIU Yi, et al. CRNet: class relation network for crop remote sensing image semantic segmentation[J]. Journal of Image and Graphics, 2022, 27(11): 3382 - 3394. (in Chinese)

[10] 施玮囡. 基于深度学习的植物工厂幼苗表型信息获取及定植装备研究[D]. 杭州:浙江大学,2019.
SHI Weinan. Study on deep learning based seedling phenotyping extraction and automatic transplanter in plant factory[D]. Hangzhou: Zhejiang University, 2019. (in Chinese)

[11] KOLODIAZHNYI M, VORONTSOVA A, KONUSHIN A, et al. Oneformer3d: one transformer for unified point cloud segmentation[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 20943 - 20953.

[12] XIANG L, BAO Y, TANG L, et al. Automated morphological traits extraction for sorghum plants via 3D point cloud data analysis[J]. Computers and Electronics in Agriculture, 2019, 162: 951 - 961.

[13] QI C R, SU H, MO K, et al. Pointnet: deep learning on point sets for 3D classification and segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 652 - 660.

[14] LI D, LI J, XIANG S, et al. PSegNet: simultaneous semantic and instance segmentation for point clouds of plants[J]. Plant Phenomics, 2022, 2022: 9787643.

[15] LI D, SHI G, LI J, et al. PlantNet: a dual-function point cloud segmentation network for multiple plant species[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2022, 184: 243 - 263.

[16] 王纪章,姚承志,周静,等. 基于三维点云的黄瓜叶片分割与表型参数提取方法[J]. 农业机械学报, 2025,56(3):354 - 362.
WANG Jizhang, YAO Chengzhi, ZHOU Jing, et al. Cucumber leaf segmentation and phenotypie extraction method based on three-dimensional point cloud[J]. Transations of the Chinese Society for Agricultural Machinery, 2025,56(3):354 - 362. (in Chinese)

[17] KERBL B, KOPANAS G, LEIMKÜHLER T, et al. 3D Gaussian splatting for real-time radiance field rendering[J]. ACM Trans. Graph., 2023, 42(4): 1 - 14.

[18] 李亚光,李洁,薛廷玉,等. 基于 PSNR 和 SSIM 方法评估双能量 CT 肝脏虚拟平扫图像质量研究[J]. CT 理论与应用研究(中英文),2025,34(1):51 - 57.

LI Yaguang, LI Jie, XUE Tingyu, et al. Evaluation of dual energy CT liver virtual plain scan image quality based on PSNR and SSIM methods[J]. CT Theory and Application Research, 2025, 34(1): 51－57. (in Chinese)

[19] YANG X, FU T, DAI G, et al. An efficient accelerator for point-based and voxel-based point cloud neural networks[C]// 2023 60th ACM/IEEE Design Automation Conference (DAC). IEEE, 2023: 1－6.

[20] QI C R, YI L, SU H, et al. Pointnet++: deep hierarchical feature learning on point sets in a metric space[J]. arXiv:1706.02413,2017.

[21] GUO Y, WANG H, HU Q, et al. Deep learning for 3D point clouds: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43(12): 4338－4364.

[22] HUANG Q, WANG W, NEUMANN U. Recurrent slice networks for 3D segmentation of point clouds[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 2626－2635.

[23] WU Z, SONG S, KHOSLA A, et al. 3D shapenets: a deep representation for volumetric shapes[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1912－1920.

[24] FAN H, SU H, GUIBAS L J. A point set generation network for 3D object reconstruction from a single image[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 605－613.

[25] HUTTENLOCHER D P, KLANDERMAN G A, RUCKLIDGE W J. Comparing images using the Hausdorff distance[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(9): 850－863.

[26] KULLBACK S. Information theory and statistics[M]. Courier Corporation, 1997.

[27] RETHAGE D, WALD J, STURM J, et al. Fully-convolutional point networks for large-scale point clouds[C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 596－611.

[28] 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, 2016: 770－778.

[29] ZHU K, WU J. Residual attention: a simple but effective method for multi-label recognition[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 184－193.

[30] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998－6008.

[31] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2020: 12993－13000.

[32] GUO M H, CAI J X, LIU Z N, et al. PCT: point cloud transformer[J]. Computational Visual Media, 2021, 7: 187－199.

[33] XU M, DING R, ZHAO H, et al. Paconv: position adaptive convolution with dynamic kernel assembling on point clouds[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 3173－3182.

[34] XU M, ZHANG J, ZHOU Z, et al. Learning geometry-disentangled representation for complementary understanding of 3D object point cloud[C]// Proceedings of the AAAI Conference on Artificial Intelligence, 2021: 3056－3064.

[35] DENG D. DBSCAN clustering algorithm based on density[C]// 2020 7th International Forum on Electrical Engineering and Automation (IFEEA). IEEE, 2020: 949－953.

[36] XIANG L, BAO Y, TANG L, et al. Automated morphological traits extraction for sorghum plants via 3D point cloud data analysis[J]. Computers and Electronics in Agriculture, 2019, 162: 951－961.