Title:

Segmentation of Plant-part from 3D Point Cloud Using Deep Learning and Multi-view Vision

Authors:
Yibin Lai, laiyibee@163.com, Guangxi Normal University
Shenglian Lu, shllu@126.com, Guangxi Normal University
Tingting Qian, qiantingting@saas.sh.cn, Shanghai Academy of Agricultural Sciences
Ming Chen, mingchen@mailbox.gxnu.edu.cn, Guangxi Normal University
Zhen Song, 443671502@qq.com, Guangxi Normal University
Guo Li, liguo@mailbox.gxnu.edu.cn, Guangxi Normal University

Keywords:
Deep learning, Point cloud, Plant phenotyping, 3D-BoNet, Instance segmentation, Multi-view vision

DOI: 10.14733/cadconfP.2021.218-222

Introduction:

Automatic measuring and monitoring technologies make great revolution to modern agriculture in the 21st century. High-throughput, precision, and non-destructive plant phenotyping measurement has become a key issue in agricultural field. Researchers obtain 3D structure information of plant by reconstructing 3D models of plant such as point cloud [5, 6]. Point cloud is a kind of data that can represent 3D depth information. In recent years, the application of deep learning methods on images has achieved amazing results. However, the application of deep learning methods on point cloud is still potential for exploration.

The existing deep learning methods for point cloud data [4] are mainly developed from two aspects. One is non-point-based methods , such as multi-view-based [9], and voxel-based [12]. The other is point-based method, such as PointNet [8] network, which provides an end-to-end learning method that can directly process point cloud. It is strong, but it does not have the ability to capture local information, so PointNet++ [7] was proposed for iterative feature extraction through the field of each point so that the network can better extract the local features of the point cloud.

In plant phenotyping research, we want to automatically obtain plant phenotypic parameters through three-dimensional point clouds. One of the key steps is to segment the point clouds of different scenes by an instance. Then we can obtain the number, position and size of the plants in the scene. We noticed that the 3D-BoNet [11] point cloud instance segmentation network has the characteristics of simple design, universal and efficient. The above network characteristics are worthy that we research deep learning application for the acquisition of plant phenotypic parameters based on this framework.

In this study, we consider two segmentation tasks from 3D point cloud:

* Segmentation of plant from group scene: We have tried place different plant species, numbers, and separation distance to obtain different plant group scene data. In order to meet the needs of monitoring the growth of plants, it is necessary for us to automatically calculate how many plants exist in this scene.

Proceedings of CAD'21, Barcelona, Spain, July 5-7, 2021, 218-222

* Segmentation of organs from an individual plant: We not only separate the ground from the plants, but also to separate different individual plants in this plant group. So this task is to segment each organ(such as leaf, fruit) from the point cloud of an individual plant, this segmentation would further be processed for extracting some phenotypic parameters, such as leaf length, leaf width, the number of each category organ and so on.

Plant Datasets Collection and Annotation:

Three crops were chosen for collecting images, including cucumbers, eggplants and tomatoes. The multi-view images of each plant group or individual include 40 to 50 images shot from different positions distributed in a hemispherical shape around the target plant. As shown in Fig.1. We used the SFM-MVS (structure from motion and multi-view stereo) method to reconstruct the 3d point cloud. The collected images were input into VisualSFM [2] software to automatically reconstruct the 3D point cloud of plant population, the processing as shown in Fig.2. And the point clouds of different plant are shown in Fig.3. Besides the point clouds of group plants generated from multi-view images, a group point clouds of individual cucumber plants are also prepared. These point clouds are obtained through parameterized 3D modeling software, as shown in the Fig.4(a). Then we performed the annotation work on these point clouds similar to the S3DIS dataset [1] to obtain data that can be trained and tested. We use Cloudcompare software to divide the points of different semantic labels and different instance labels into different parts, as shown in Fig.4.

(a) Original cucumber photo. (b) Original eggplant photo. (c) Original tomato photo.

Fig. 1: The original photos were taken using a Canon EOS 5D Mark III digital camera.

(a) The 42 multi-view photos of the same plant. (b) Sparse point cloud generated by SFM method. (c) Dense point cloud generated by MVS method.

Fig. 2: The processing of SFM-MVS method to reconstruct plant point cloud.

Point Cloud Instance Segmentation Network:

The main structure of 3D-BoNet is divided into a backbone network and two branch networks:

* **Backbone network:** The backbone network is mainly responsible for the feature extraction of

(a) Original cucumber point cloud. (b) Original eggplant point cloud. (c) Original tomato point cloud.

Fig. 3: Point cloud of different kinds of plants.

(a) individual plant point cloud. (b) Semantic label. (c) Instance label. (d) Semantic label. (e) Instance label.

Fig. 4: Annoatation of plant point cloud.

point clouds. In 3D-BoNet, PointNet or PointNet++ is used as the backbone network to extract the local and global features of the point cloud for input to the subsequent branch network.

* **The bounding box prediction branch:** The core part of 3D-BoNet is the prediction of the bounding box branch. The main purpose of this branch is to predict the bounding box of the instance through the global feature information extracted from the backbone network.

Experiments:

We train the instance and semantic segmentation network on the plant datasets based on this framework. The semantic segmentation branch is the same as used in [10]. While using the training set of different scenes, we also changed the two parts in the training process.

Training environment: Our training and testing are running on GTX 1660 super graphics card and Intel i5 CPU core under Ubuntu16.04, tensorflow1.4, python3.6 and cuda 8.0. The settings of the training datasets and testing datasets are shown in the Tab.1.

Using different backbone networks: The framework does not limit the use of any kind of network. In our training experiment, we use PointNet and PointNet++ backbone networks to train.

Data rotation augmentation based on point cloud: The leaves in our plant point cloud dataset have obvious changes in the rotation characteristics, and the angles and directions of each leaf are varied. Therefore, we need to augment the point cloud rotation of the dataset. After rotatation augumentation in different directions and different angles, these differences are used to train the network. We rotate the point cloud around the X-axis, Y-axis, Z-axis or X-Y-Z axis to create a new data set.

Results and Evaluations:

After each dataset had completed 100 epoch of training, the obtained network model is used to predict semantic and instance label on the test point cloud. The visual result of the best prediction is shown in the Fig.5 and Fig.6; When the threshold of IoU was set to 0.5 to evaluate the Mean Precision and Mean Recall, we got the results shown in Tab.2 and Tab.3.

	Number of training datasets	Number of testing datasets
Individual plant dataset)	32	4
Population plant dataset	16	4

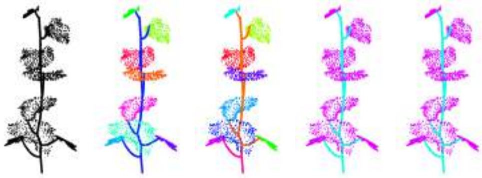
Table 1: Training datasets and testing datasets settings

(a) original point cloud (b)ins_label (c)ins_pred (d) sem_label (e)sem_pred

Fig. 5: Segmentation results of plant population point cloud.

	original dataset		Rot:XYZ	
backbone	mPrec	mRec	mPrec	mRec
PointNet	0.788	0.667	\	\
PointNet++	**0.875**	**0.897**	0.829	0.743

Table 2: The mean precision and mean recall under different backbone networks and different augmentation plant population datasets.



(a) original point cloud (b)ins_label (c)ins_pred (d) sem_label (e)sem_pred

Fig. 6: Segmentation results of individual plant point cloud.

	original dataset		rot: X		rot: X+Y+Z+XYZ	
backbone	mPrec	mRec	mPrec	mRec	mPrec	mRec
PointNet	0.543	0.442	0.571	0.55	0.674	0.721
PointNet++	0.636	0.488	0.634	0.634	**0.809**	**0.884**

Table 3: The mean precision and mean recall under different backbone networks and augmentation datasets.

Conclusions:

According to the above results and evaluation, we cand reach the following conclusions: (1) Detailed point cloud features extracted from the backbone network can improve the effect of instance segmentation; (2) Point cloud rotation augmentation has a significant impact in improving the segmentation effect of individual leaf instances. This paper explores the application of deep learning in plant point cloud segmentation, which provides a direction for the subsequent research on automatic acquisition of plant phenotypic parameters. The potential of deep learning-based point cloud processing in plant phenotyping research maintains a hot topic and open challenge, which deserves more people's attention.

Acknowledgement:
This research is supported by National Natural Science Foundation of China (No. 61762013, 62062015 and 61662006), the Science and Technology Foundation of Guangxi Province (No. 2018AD19339) and Research Fund of Guangxi Key Lab of Multi-source Information Mining & Security(No. 20-A-02-02).

References:
[1] Armeni, I.; Sener, O.; Zamir, A.; Jiang, H.; Brilakis, I.; Fischer, M.; Savarese, S.: 3d semantic parsing of large-scale indoor spaces. Computer Vision and Pattern Recognition, 2016. http://doi.org/10.1109/CVPR.2016.170

[2] Furukawa, Y.; Curless, B.; Seitz, S.M.; Szeliski, R.: Towards internet-scale multi-view stereo. Computer Vision and Pattern Recognition, 2010. http://doi.org/10.1109/CVPR.2010.5539802

[3] Geiger, A.; Lenz, P.; Urtasun, R.: Are we ready for autonomous driving? the kitti vision benchmark suite. Computer Vision and Pattern Recognition, 3354-3361, 2012. http://doi.org/10.1109/CVPR.2012.6248074

[4] Guo, Y.; Wang, H.; Hu, Q.; Liu, H.; Bennamoun, M.: Deep learning for 3d point clouds: A survey. Pattern Analysis and Machine Intelligence, PP(99), 1-1, 2020.

[5] Hui, F.; Zhu, J.; Hu, P.; Meng, L.; Zhu, B.; Guo, Y.; Li, B.; Ma, Y.: Image-based dynamic quantication and high-accuracy 3d evaluation of canopy structure of plant populations. Annals of Botany, 5(121), 1079-1088, 2018. http://doi.org/10.1093/aob/mcy016

[6] Jin, S.; Su Y.; Wu, F.; Pang, S.; Gao, S.; Hu, T.; Liu, J.; Guo, Q.: Stem-leaf segmentation and phenotypic trait extraction of individual maize using terrestrial lidar data. Geoscience and Remote Sensing, 3(57), 2019. http://doi.org/10.1109/TGRS.2018.2866056

[7] Qi, C.R.; Li, Y.; Hao, S.; Guibas, L.J.: Pointnet++: Deep hierarchical feature learning on point sets in a metric space. Neural Information Processing Systems, 2017.

[8] Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J.: Pointnet: Deep learning on point sets for 3d classication and segmentation. Computer Vision and Pattern Recognition, 77-85, 2017. http://doi.org/10.1109/CVPR.2017.16

[9] Su, H.; Maji, S.; Kalogerakis, E.; Learned-Miller, E.: Multi-view convolutional neural networks for 3d shape recognition. International Conference on Computer Vision, 945-953, 2015. http://doi.org/10.1109/ICCV.2015.114

[10] Wang, W.; Yu, R.; Huang, Q.; Neumann, U.: Sgpn: Similarity group proposal network for 3d point cloud instance segmentation. Computer Vision and Pattern Recognition, 2017. http://doi.org/10.1109/CVPR.2018.00272

[11] Yang, B.; Wang, J.; Clark, R.; Hu, Q.; Wang, S.; Markham, A.; Trigoni, N.: Learning object bounding boxes for 3d instance segmentation on point clouds. Neural Information Processing Systems, 2019.

[12] Yin, Z.; Tuzel, O.: Voxelnet: End-to-end learning for point cloud based 3d object detection. Computer Vision and Pattern Recognition, 2017.