

# An End-to-End Deep RNN based Network Structure to Precisely Regress the Height of Lettuce by Single Perspective Sparse Point Cloud

Jinsong Li, Xiyue Guo, Ying Wang, Maowei Li, Lihua Zheng, Minjuan Wang\*  
Key Laboratory of Modern Precision Agriculture System Integration Research,  
Ministry of Education, China Agricultural University, China  
{minjuan, jingsong, zhenglh, sy20203081578}@cau.edu.cn leowyztl@163.com

## 1. Introduction

The point cloud data contains real three-dimensional coordinate information of the object. Compared with ranging by two-dimensional image, in theory, point cloud ranging needs no various perspective corrections or special shooting conditions sampling due to the impact of shooting angle of view and shooting distance. Is it an excellent idea to automatically obtain the three-dimensional phenotype of lettuce by processing the point cloud data of lettuce through an end-to-end training network?

In recent years, researchers have achieved amazing achievements on point cloud data sets using deep learning networks. These achievements are sufficient to prove the superiority of deep learning networks in point cloud feature extraction. However, most of these breakthrough studies focused on single classification, multiple classification and segmentation tasks [3, 4, 5]. In addition to classification tasks, plant phenotype research also involves a large number of regression tasks, such as plant height, biomass, and leaf area, etc [6, 1]. This research aims to use lettuce point cloud data to regress its plant height, and has proposed an end-to-end deep learning network structure named DRN that can predict lettuce plant height more effectively in Figure 1(b). This structure has shown good results in the two classic point cloud feature extraction basic networks, PointNet++ and PointCNN.

## 2. Proposed Method

The DRN structure requires a basic network to extract features. Take DRNPointCNN as an example, the 1024 dimension sparse point cloud(batchsize $\times$ 1024 $\times$ 3) were processed by 4-layer Xconv and a global average pooling in the PointCNN network to obtain 384 dimension embeddings. After that, input these embedding into BiLSTM with input size of 1 and output size of 1 to learn the relationship between embeddings. From BiLSTM we got a new embedding with the dimension of 2 $\times$ 384, then we used pooling layer to deal with these new embeddings, and sent the re-

sults to a regression network to predict plant height.

The experimental data came from the Third Greenhouse Growing Challenge [2]. The images were captured by RealSense camera and contain data of individual lettuce plants of different species in different growth period. Each sample is connected with the ground truth plant height from manual measurement. We transformed these 1920 $\times$ 1080 resolution images and corresponding depth value into cloud point efficiently using Matrix Operations and CUDA acceleration, and then obtained the sparse cloud points through surface downsampling and filtering in Figure 1(a).

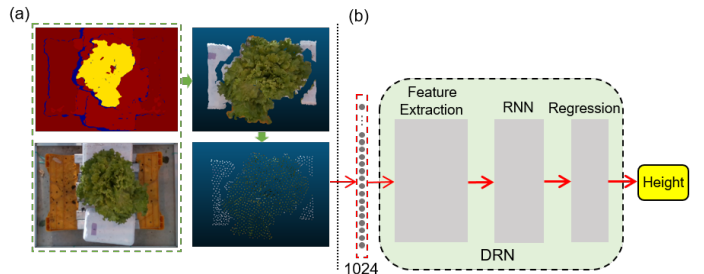


Figure 1. Overview of our Method.(a)Data and Data preprocessing, transform the give images into point cloud data and downsampling. (b) Regressing lettuce plant height from single pespective point cloud of 1024 dimension by proposed DRN structure model.

## 3. Result and Discussion

A total of 341 lettuce point cloud samples of 4 types in different periods were used for the experiment, of which 301 samples were selected as training data and 40 samples were used as testing data. We used MSE as the loss function of the model, simultaneously, the MSE value between the predicted value and ground truth were used as the evaluation index of the accuracy, the lower the MSE, the more accurate the prediction of the lettuce plant height by the model.

Figure 2 represents MSE result in different epoch on the training data from RPNPointCNN based Regression model

Table 1. Performance Analysis of DRN Structure in PointCNN and PointNet++

Methods	PointCNN	DRNPointCNN	PointNet++	DRNPointNet++
Train MSE (cm)	1.72	<b>1.27</b>	<b>2.06</b>	2.11
Test MSE (cm)	3.74	<b>2.86</b>	3.45	<b>2.16</b>

and PointCNN based Regression model. It can be clearly seen that DRNPointCNN has much lower MSE compared to PointCNN on training data, which means DRNPointCNN has better fitting ability. Correspondingly, the MSE curve of the testing data in Figure 3 can also reflect DRNPointCNN is supposed to give a more accurate result of lettuce plant height. Figure 4 and Figure 5 show the performance of DRN in another classic point cloud feature extraction network. Figure 4 shows that the MSE of DRNPointNet++ on training data does not reflect its effectiveness, while the curve in Figure 5 shows DRNPointNet++ has better performance than PointNet++ on testing data. In order to further explore the performance difference of the above models on testing data, we selected each trained model weight at 500th epoch to predict the lettuce plant height for 50 times and got the average value to calculate MSE. The prediction results are shown in Table 1. Compared with PointCNN, the DRNPointCNN base regression model yields a smaller MSE value on both testing data and training data, while DRNPointNet++ decreases the MSE value on testing data.

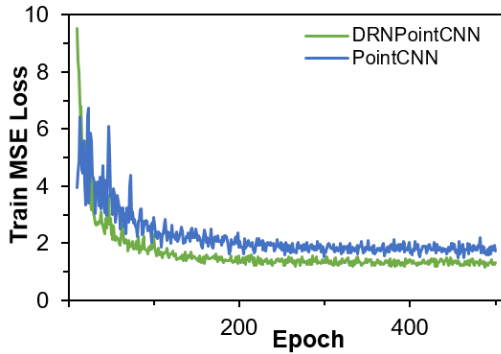


Figure 2. MSE of DRNPointCNN and PointCNN on training data.

#### 4. Conclusion

In this research, we present a Deep RNN Based Network Structure for point cloud regression task called DRN. It has been proven to achieve accuracy improvements in PointNet++ and PointCNN when it comes to regression of lettuce plant height. On the testing data, the MSE of DRNPointNet++ is 37.39% lower than that of PointNet++, and the MSE of DRNPointCNN is 23.53% lower than PointCNN.

We believe DRN structure is suitable for feature extraction from single-view sparse plant point cloud data along with regression of spatial distance related plant phenotypes

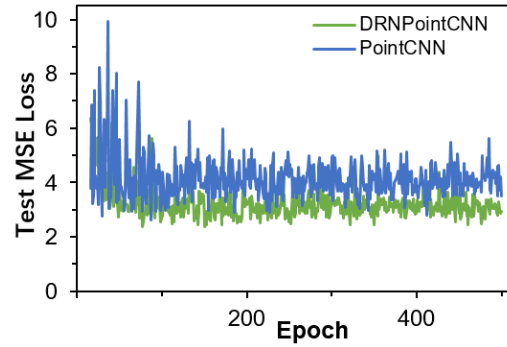


Figure 3. MSE of DRNPointCNN and PointCNN on testing data.

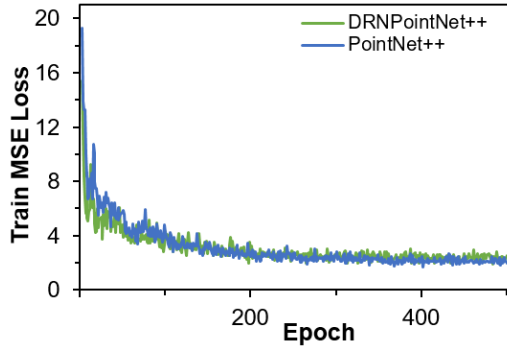


Figure 4. MSE of DRNPointNet++ and PointNet++ on training data.

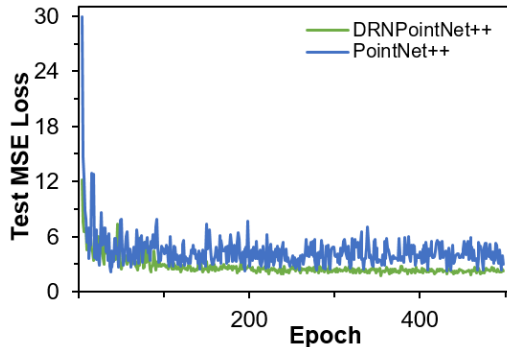


Figure 5. MSE of DRNPointNet++ and PointNet++ on testing data.

like plant height. The operator we used to build model are mainly from pytorch and pytorch-geometric. The source code and datasets of our research can be found at <https://github.com/LeJson/DRN>, we hope it can encourage further development in plant phenotype research based on point cloud data.

## References

- [1] Jan U.H. Eitel, Troy S. Magney, Lee A. Vierling, Tabitha T. Brown, and David R. Huggins. Lidar based biomass and crop nitrogen estimates for rapid, non-destructive assessment of wheat nitrogen status. *Field Crops Research*, 159:21–32, 2014.
- [2] S. (Silke) Hemming, H.F. (Feije) de Zwart, A. (Anne) Elings, monique bijlaard, van Bart Marrewijk, and Anna Petropoulou. 3rd Autonomous Greenhouse Challenge: Online Challenge Lettuce Images. 8 2021.
- [3] Yangyan Li, Rui Bu, Mingchao Sun, Wei Wu, Xinhan Di, and Baoquan Chen. Pointcnn: Convolution on x-transformed points. In *NeurIPS*, pages 828–838, 2018.
- [4] Charles Ruizhongtai Qi, Li Yi, Hao Su, and Leonidas J. Guibas. Pointnet++: Deep hierarchical feature learning on point sets in a metric space. In *NIPS*, pages 5105–5114, 2017.
- [5] Hugues Thomas, Charles R. Qi, Jean-Emmanuel Deschaud, Beatriz Marcotegui, Francois Goulette, and Leonidas J. Guibas. Kpconv: Flexible and deformable convolution for point clouds. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 8 2019.
- [6] Jian Zhang, Tianjin Xie, Wanneng Yang, and Guangsheng Zhou. Crop height estimation based on uav images: Methods, errors, and strategies. *Computers and Electronics in Agriculture*, 185:106155, 2021.