Robustness of 3D point-based deep learning for plant organ segmentation against point density variation and noise

David Rousseau[1], Kaya Turgut[2], and Helin Dutagaci[2]

[1]Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS), Université d'Angers, UMR IRHS, INRAe
[2]Department of Electrical-Electronics Engineering, Osmangazi University

October 5, 2022

Robustness of 3D point-based deep learning for plant organ segmentation against point density variation and noise

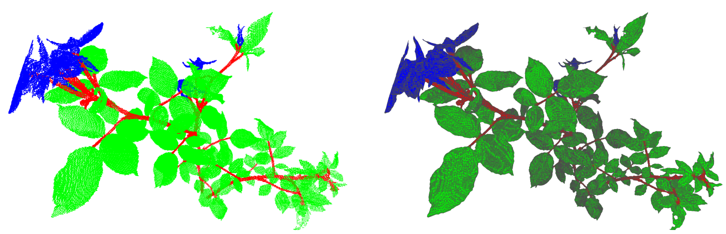
Kaya Turgut[a], Helin Dutagaci[a], and David Rousseau[b]

[a]Department of Electrical-Electronics Engineering, Osmangazi University, Eskisehir, Turkey
[b]Université d'Angers, Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS), UMR IRHS, INRAe, Angers, France

ABSTRACT

We investigate the robustness of 3D point-based deep learning for organ segmentation of 3D plant models against varying reconstruction quality of the surface. The reconstruction quality is quantified in two ways: 1) The number of acquisitions for partial 3D scans and 2) the amount of noise. High quality models of real rosebush plants are used to collect point clouds in a controlled simulation environment as a way to degrade surface quality systematically. We show that the well-known 3D point-based neural network PointNet++ is capable of operating effectively on low quality and corrupted data for the task of plant organ segmentation. The results indicate that investing on developing deep learning methods has the potential of advancing applications of automated phenotyping, especially for low-quality 3D point clouds of plants.

Keywords: plant phenotyping, organ segmentation, robustness analysis, point-based deep learning



(a) (b)

Figure 1: A 3D rosebush model from ROSE-X data set: (a) point cloud; (b) triangular mesh model.

1. INTRODUCTION

One common representation of the 3D geometry of plants is the point cloud, which can be collected directly from depth scanners such as LiDARs[1,2] and RGB-D cameras[3] or indirectly from 2D images through multi-view 3D reconstruction.[4,5] Data acquisition with such devices should be performed at various viewpoints to cover the entire plant for reconstruction of a complete 3D model. The number of viewing points is an important parameter for a standardized data acquisition protocol, heavily impacting the quality of the reconstructed surface and acquisition time and/or complexity. The quality of the reconstructed surface also depends on the noise characteristics of the sensor. The analysis of the behaviour of computational tools for extraction of 3D geometrical traits under the variation of reconstruction noise and number of views is important for sensor choice and determination of the data acquisition protocol.

In the previous works, the impact of noise was assessed under different level of Gaussian noise[6–8] and the robustness analyses were performed against surface quality[9,10] for classical supervised or unsupervised methods. Works in the similar vein are needed to evaluate robustness of 3D point-based deep learning methods for plant phenotyping tasks. Employment of 3D point-based neural networks for plant phenotyping has just very recently

Further author information: (Send correspondence to David Rousseau)
David Rousseau: E-mail: david.rousseau@univ-angers.fr, ORCID: xxxx-xxxx-xxxx-xxxx

emerged, and has been demonstrated to be effective on point clouds acquired by various imaging modalities.[10–14] To the best of our knowledge, the only work investigating the effect of surface quality on the segmentation performance of 3D deep learning methods is conducted by Chaudhury et al.[15] They assessed two sampling strategies on PointNet++[16] on virtual plants of Arabidopsis thalian. Inspired by the suggestions in Ref. 17 to model characteristics of real world scanning devices, we simulated various degrees of reconstruction quality in terms of point density, self-occlusion, and noise.

In this paper, we provide such analysis for plant organ segmentation on 3D point-based deep learning, specifically on PointNet++.[16] Moreover, the comparison with a traditional machine learning technique based on hand-crafted local surface features and Support Vector Machines (SVM) is performed. In our previous work,[18] six point-based deep learning networks were compared for plant segmentation, with PointNet++ yielding the highest performance. However, the 3D plant models used in the previous study were of high quality and relatively noise-free. The purpose of this study is to supply a systematic analysis of the robustness of PointNet++ against degrading reconstruction quality.

2. MATERIALS AND METHODS

Data set: The publicly available ROSE-X data set[19] was used for experimental evaluation. The data set consists of 11 complete and fully annotated 3D models of real rosebush plants. Each point in the 3D point cloud was manually assigned to one of flower, leaf, and stem categories. In Fig. 1, a 3D rosebush model from ROSE-X data set is illustrated. For the simulation of data acquisition through a 3D sensor, a ToF camera in our case, the point clouds are converted to triangular mesh models through Delaunay triangulation.

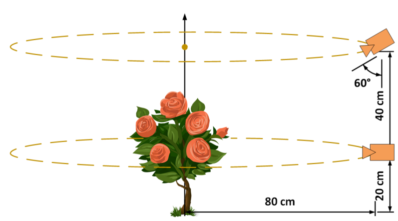
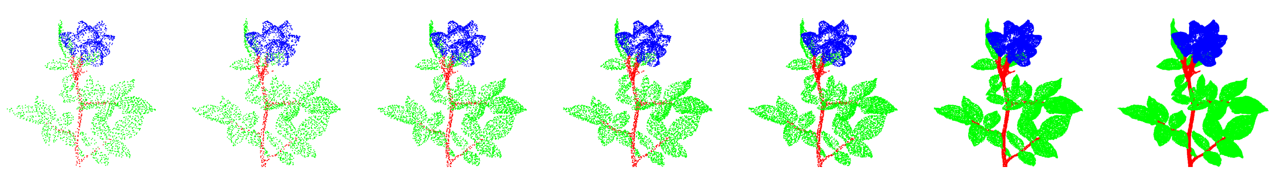


Figure 2: The acquisition system for the reconstruction of rosebush plants.

Simulation of 3D Acquisition and Reconstruction: The open-source Blender Sensor Simulation plugin Blensor[20] was used for the simulation of a ToF camera. The 3D scanning process is depicted in Fig. 2. The origin is set to the base of the rosebush. The vertical axis corresponds to the upward direction of the rosebush model. The viewpoints are restricted to two circles of radius $80cm$ around the vertical axis. One circle is located at $20cm$ height. The viewing angles are oriented perpendicular towards the vertical axis. The other circle is located at $60cm$ height. The viewing angles on this circle are oriented towards the vertical axis tilted with $60°$. We refer to these viewpoints on the two circles as (i) side-views and (ii) top-diagonal-views, respectively.



(a) (b) (c) (d) (e) (f) (g)

Figure 3: Registration outputs for seven simulations for a sample rosebush model. The viewpoints are equally spaced on the view circles at angles of multiples of (a) 120°; (b) 90°; (c) 60°; (d) 45°; (e) 30°; (f) 10°; (g) 5°.

For each of the 11 triangular mesh models in the ROSE-X data set, ToF data acquisition was simulated with varying number of viewpoints and seven 3D point clouds per model was reconstructed. The viewpoints are located on the two circles at equally spaced angles; specifically at multiples of 120°, 90°, 60°, 45°, 30°, 10°, and

5°. The number of viewpoints, hence the number of partial scans, correspond to 6, 8, 12, 16, 24, 72, and 144, respectively, combining the side-views and top-diagonal-views. Since the camera locations and orientations are known, the partial scans were registered by the simulation environment. In Fig. 3, the registration outputs of the seven simulations for a sample rosebush model are given. More viewpoints mean more coverage of the plant, less self-occlusion and higher point density. However, an imaging protocol with more viewpoints requires longer time and computational cost for acquisition, registration, and storage.

Simulation of Noise: For simulating sensor noise and reconstruction errors, Gaussian noise of varying standard deviation was applied as it is the common practice for noise modeling of 3D sensors.[21,22] Let the noise-free point on the surface as acquired by the camera at a particular view be d_i representing (x_i, y_i, z_i) coordinates in the camera reference frame measured in meters. The perturbed point p_i is obtained by $p_i = d_i + \mathcal{N}(\mu, \sigma)\frac{d_i}{\|d_i\|}$ where $\mathcal{N}(\mu, \sigma)$ is a Gaussian random variable with mean $\mu = 0$ and standard deviation σ. The point is moved along the orientation towards (or away from) the origin of the camera. The amount of the movement is determined by the randomly chosen number $\mathcal{N}(\mu, \sigma)$. The perturbed point is then represented with respect to a global reference frame for registration. The σ parameter is varied as $0.5mm$, $1mm$, $1.5mm$, $2mm$, $2.5mm$, $3mm$, and $3.5mm$ to systematically increase the amount of the noise. Surface points on a sample rosebush region without noise and perturbed with varying amounts of noise are given in Fig. 4.

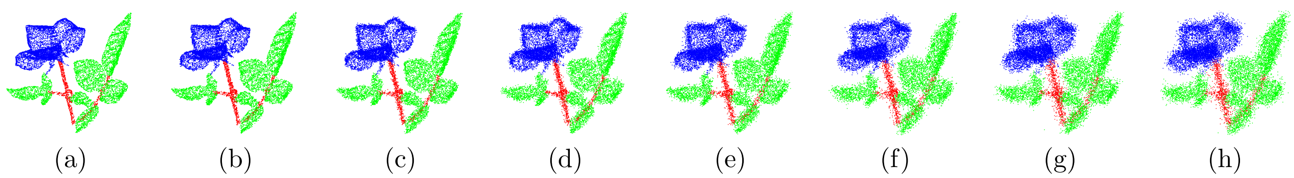


Figure 4: The perturbation of points for varying amounts of noise (σ): (a) Noise-free; (b) $0.5mm$; (c) $1mm$; (d) $1.5mm$; (e) $2mm$; (f) $2.5mm$; (g) $3mm$; (h) $3.5mm$.

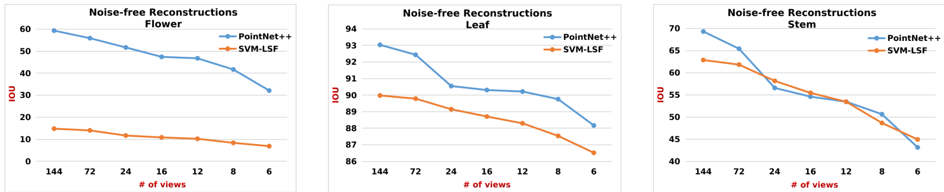
Semantic Segmentation: For segmentation of 3D plant models, the traditional approach was to use machine learning techniques together with hand-crafted local surface features.[19,23–26] Recently, adopting and developing 3D point-based neural networks to perform 3D phenotyping tasks proved to be effective.[11–15,18,27,28] Such approaches should be evaluated in terms of their robustness to point density and noise in comparison with traditional machine learning techniques. In this study, such a comparison is done between PointNet++[16] and a machine learning approach based on SVM and hand-crafted local surface features (SVM-LSF).[19]

3. RESULTS

We performed a 5-fold cross validation procedure to evaluate the segmentation performance of PointNet++ and SVM-LSF. One rosebush model is used for each fold for training the models and remaining ten rosebush models of ROSE-X are reserved as the test data set. To compare the segmentation performance of trained models, Intersection over Union (IoU) of each class and mean of these values ($MIoU$) are used.

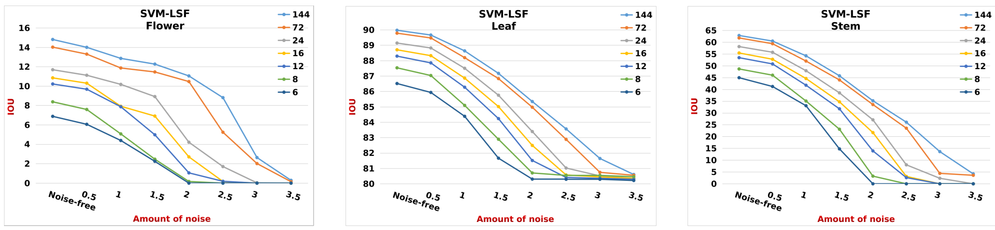
We conducted experiments on virtually reconstructed point clouds through ToF simulation, and virtually reconstructed point clouds corrupted by noise. For each reconstruction process of the ToF simulation, the partial scans acquired from angles at multiples of 120°, 90°, 60°, 45°, 30°, 10°, and 5° on the two view circles (Fig. 2) are registered. To create noisy clouds, Gaussian noise of varying amounts (measured as σ in mm) is added to partial scans before registration.

The results of SVM-LSF and PointNet++ for noise-free point clouds are given in Fig. 5. The results were averaged over the 5-fold experiments. The performance of both methods on simulated reconstructions drop as the number of scans decreases. The IoU for the flower and leaf classes obtained by PointNet++ remains much higher as compared to SVM-LSF as the point density decreases. For both methods, the performance on the stem class has the steepest drop, since the thin branches are occluded and represented by fewer points (Fig. 3). The effect is particularly dramatic for the petioles, which are classified as 'stem' in the ground-truth. Nevertheless, a 10% drop in the $mIoU$ as the number of scans decreases from 144 to 12 is demonstrative of the robustness of PointNet++ on relatively sparse point clouds.

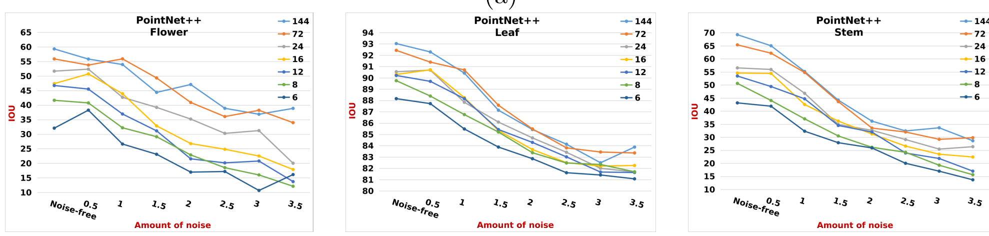


(a) (b) (c)

Figure 5: The IoU metric for SVM-LSF and PointNet++ of noise-free reconstructions with different number of views.



(a)



(b)

Figure 6: The segmentation performance in IoU with varying number of views and noise ratio (σ in mm): (a) SVM-LSF; (b) PointNet++.

The effects of increasing noise imposed over the reconstructions with varying number of scans are given in Fig. 6. PointNet++ outperformed SVM-LSF for all cases. PointNet++ continues to distinguish classes with reduced performance, while SVM starts to classify all points as leaves in response to increased noise ratio. The hand-crafted local features of SVM-LSF become less descriptive and less reliable as the surface points become more corrupted with noise. The IoUs for flower and stem classes become zero when the standard deviation of the noise reaches 2mm. PointNet++, on the other hand, continues to learn from noisy data since the contextual information at larger local scales is still present allowing the network to discern various structures (Fig. 4).

4. CONCLUSION

In this work, we presented a systematic analysis of the robustness of point-based deep learning for semantic segmentation of 3D plant models against degrading surface quality. PointNet++ and SVM-LSF were used as representatives and the surface quality was quantified in terms of number of partial scans from different viewpoints and the amount of noise. We demonstrated that the degradation of segmentation performance of PointNet++ on sparse and occluded point clouds is comparable to SVM-LSF. Furthermore, PointNet++ proved to be more robust to noise.

ACKNOWLEDGMENTS

The authors acknowledge the support of The Scientific and Technological Research Council of Turkey (TUBITAK), Project No: 121E088. The authors also acknowledge the support of 2214/A International Doctoral Research Fellowship Program granted to Kaya Turgut by TUBITAK and PHENOTIC platform node of the french plant phenotyping infrastructure PHENOME.

REFERENCES

[1] Thapa, S., Zhu, F., Walia, H., Yu, H., and Ge, Y., "A novel lidar-based instrument for high-throughput, 3d measurement of morphological traits in maize and sorghum," *Sensors* **18**(4) (2018).

[2] Chaivivatrakul, S., Tang, L., Dailey, M. N., and Nakarmi, A. D., "Automatic morphological trait characterization for corn plants via 3D holographic reconstruction," *Computers and Electronics in Agriculture* **109**, 109–123 (2014).

[3] Hu, Y., Wang, L., Xiang, L., Wu, Q., and Jiang, H., "Automatic non-destructive growth measurement of leafy vegetables based on Kinect," *Sensors* **18**(3) (2018).

[4] Nguyen, T. T., Slaughter, D. C., Max, N., Maloof, J. N., and Sinha, N., "Structured light-based 3D reconstruction system for plants," *Sensors* **15**(8), 18587–18612 (2015).

[5] Wang, Y., Wen, W., Wu, S., Wang, C., Yu, Z., Guo, X., and Zhao, C., "Maize plant phenotyping: Comparing 3d laser scanning, multi-view stereo reconstruction, and 3d digitizing estimates," *Remote Sensing* **11**(1) (2019).

[6] Malambo, L., Popescu, S., Horne, D., Pugh, N., and Rooney, W., "Automated detection and measurement of individual sorghum panicles using density-based clustering of terrestrial lidar data," *ISPRS Journal of Photogrammetry and Remote Sensing* **149**, 1–13 (2019).

[7] Hämmerle, M. and Höfle, B., "Effects of reduced terrestrial LiDAR point density on high-resolution grain crop surface models in precision agriculture," *Sensors* **14**(12), 24212–24230 (2014).

[8] Sodhi, P., Sun, H., Póczos, B., and Wettergreen, D., "Robust plant phenotyping via model-based optimization," in [*2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*], 7689–7696 (2018).

[9] Ando, R., Ozasa, Y., and Guo, W., "Robust surface reconstruction of plant leaves from 3D point clouds," *Plant Phenomics (Washington, D.C.)* **2021**, 3184185 (2021).

[10] Ghahremani, M., Williams, K., Corke, F., Tiddeman, B., Liu, Y., Wang, X., and Doonan, J. H., "Direct and accurate feature extraction from 3D point clouds of plants using RANSAC," *Computers and Electronics in Agriculture* **187**, 106240 (2021).

[11] Li, Y., Wen, W., Miao, T., Wu, S., Yu, Z., Wang, X., Guo, X., and Zhao, C., "Automatic organ-level point cloud segmentation of maize shoots by integrating high-throughput data acquisition and deep learning," *Computers and Electronics in Agriculture* **193**, 106702 (2022).

[12] Wang, L., Zheng, L., and Wang, M., "3D point cloud instance segmentation of lettuce based on Part-Net," in [*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*], 1647–1655 (June 2022).

[13] Morel, J., Bac, A., and Kanai, T., "Segmentation of unbalanced and in-homogeneous point clouds and its application to 3D scanned trees," *The Visual Computer (Special issue for CGI 2020)* **36**, 2419–2431 (2020).

[14] Boogaard, F. P., van Henten, E. J., and Kootstra, G., "Boosting plant-part segmentation of cucumber plants by enriching incomplete 3D point clouds with spectral data," *Biosystems Engineering* **211**, 167–182 (2021).

[15] Chaudhury, A., Hanappe, P., Azaïs, R., Godin, C., and Colliaux, D., "Transferring PointNet++ Segmentation from Virtual to Real Plants," in [*ICCV 2021 - International Conference on Computer Vision*], 1–3 (Oct. 2021).

[16] Qi, C. R., Yi, L., Su, H., and Guibas, L. J., "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," in [*Advances in Neural Information Processing Systems*], Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., eds., **30**, Curran Associates, Inc. (2017).

[17] Chaudhury, A., Boudon, F., and Godin, C., "3D plant phenotyping: All you need is labelled point cloud data," in [*CVPPP-ECCV 2020 - Workshop on Computer Vision Problems in Plant Phenotyping*], 1–17 (Aug. 2020).

[18] Turgut, K., Dutagaci, H., Galopin, G., and Rousseau, D., "Segmentation of structural parts of rosebush plants with 3D point-based deep learning methods," *Plant Methods* **18**, 00857–3 (Feb. 2022).

[19] Dutagaci, H., Rasti, P., Galopin, G., and Rousseau, D., "ROSE-X: An annotated data set for evaluation of 3D plant organ segmentation methods," *Plant Methods* **16** (2020).

[20] Gschwandtner, M., Kwitt, R., Uhl, A., and Pree, W., "BlenSor: Blender sensor simulation toolbox," in [*Proceedings of the 7th International Conference on Advances in Visual Computing - Volume Part II*], *ISVC'11*, 199–208, Springer-Verlag, Berlin, Heidelberg (2011).

[21] Fang, J., Zhou, D., Yan, F., Zhao, T., Zhang, F., Ma, Y., Wang, L., and Yang, R., "Augmented LiDAR simulator for autonomous driving," *IEEE Robotics and Automation Letters* **5**(2), 1931–1938 (2020).

[22] Griffiths, D. and Boehm, J., "Synthcity: A large scale synthetic point cloud," (2019).

[23] Elnashef, B., Filin, S., and Lati, R. N., "Tensor-based classification and segmentation of three-dimensional point clouds for organ-level plant phenotyping and growth analysis," *Computers and Electronics in Agriculture* **156**, 51–61 (2019).

[24] Wahabzada, M., Paulus, S., Kersting, K., and Mahlein, A.-K., "Automated interpretation of 3D laserscanned point clouds for plant organ segmentation," *BMC Bioinformatics* **16** (2015).

[25] Sodhi, P., Vijayarangan, S., and Wettergreen, D., "In-field segmentation and identification of plant structures using 3D imaging," in [*2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*], 5180–5187 (2017).

[26] Dey, D., Mummert, L., and Sukthankar, R., "Classification of plant structures from uncalibrated image sequences," in [*2012 IEEE Workshop on the Applications of Computer Vision (WACV)*], 329–336 (2012).

[27] Boogaard, F. P., van Henten, E. J., and Kootstra, G., "Improved point-cloud segmentation for plant phenotyping through class-dependent sampling of training data to battle class imbalance," *Frontiers in Plant Science* **13**, 838190 (2022).

[28] Turgut, K., Dutagaci, H., and Rousseau, D., "Rosesegnet: An attention-based deep learning architecture for organ segmentation of plants," *Biosystems Engineering* **221**, 138–153 (2022).