Optimal plant part segmentation using 3D neural architecture search

Farah Saeed[a], Changying "Charlie" Li[a,b], Tianming Liu[a]
[a]School of Computing, University of Georgia, Athens, GA, USA; [b]College of Engineering, University of Georgia, Athens, GA, USA

ORCiD: 0000-0002-5881-722X

ABSTRACT

The automatic, and accurate plant phenotyping plays important role to improve the crop yield through enabling efficient plant analysis and plant breeding studies. The 3d deep learning has allows automatic segmentation of plant parts from point cloud data. However, the network architecture is designed manually and performance is limited to prior experience. The aim of this study is to search for optimal 3d deep networks to perform the plant part segmentation. We perform the 3d neural architecture search by training a super network composed of candidate networks. Using the trained super network, the evolutionary searching is used to search for top performing architecture. The results demonstrate the searched architecture outperforms manually designed architectures by attaining mean IoU and accuracy of more than 90% and 96%, respectively. The searched architecture achieves more than 83% class-wise IoU for all main stem, branches, and boll class. These plant part segmentation method shows promising results and holds potential to be utilized by plant breeders for enhancing the production quality.

Keywords: Plant phenotyping, Plant part segmentation, LiDAR, Point cloud, 3D Deep learning, 3D Neural architecture search

1. INTRODUCTION

Automatic plant part segmentation is essential to analyze plant organs and enhance crop yield and production quality. In addition to reducing manual labor and time, it allows for non-destructive analysis. With recent advancements in remote sensing, both 2D and 3D devices have shown progress in plant part segmentation. The 2d devices allow collecting the plants data from a single view in the form of an rgb image. However, the data in the image may be occluded. On the other hand, the data collected using the 3d sensors contains the depth information. It offers more accurate estimation of plants phenotypic traits.

Plant part segmentation using traditional processing methods have been performed for a wide range of plants, including sorghum [1], maize [2, 3], cotton [4], and others using processing methods of clustering, region growth, skeleton extraction. Parts of tomato and rosebush plants were segmented using machine learning. In this, the local point features such as normal vectors, eigen values of covariance matrix, FPFH features were utilized to segment the plant parts through classification of each point in the plant point cloud. To automatically extract the features from the data, 3D deep learning has been utilized to segment parts from point cloud data. 3D deep networks were trained to perform segmentation of various crops and plants such as wheat, cotton, rosebush and maize. In these studies, the network architectures are manually designed, and therefore the design of network architecture depends upon outcomes from prior experiments. As a result, the best performing architecture from the search space may remain unexplored. Selecting each network from the search space and training from scratch is time consuming and infeasible. The aim of this study is to search for network architecture to achieve optimal plant part segmentation using 3D neural architecture search. The specific objectives were 1) To apply 3D neural architecture search based on super network training and evolutionary searching for plant part segmentation. 2) To compare the performance of searched network architecture with manually designed architectures in baselines.

2. MATERIALS AND METHODS

2.1 Data description and preprocessing

The LiDAR dataset for cotton plants was collected in Plant research Farm in University of Georgia. Both indoor and outdoor areas were used in scanning the plant. The FARO LiDAR device was used to scan the plants from multiple views and spherical targets were used in the registration of those views. In the collected datatset, each sample covers a point cloud for a single plant. Each plant point cloud contains x,y,z coordinates and rgb i.e. color value for each point. The collected point clouds were passed through a preprocessing stage where denoising, point-wise labelling, down-sampling, normalization, and augmentation were applied. The point clouds were denoised using statistical outlier removal method where 6 nearest neighbors were selected and mean and standard deviation of their distances were estimated. The neighbors having distances greater than 0.01 times that of standard deviation were removed to denoise the point cloud. The plant part labelling was performed to annotate the main stem, branches and cotton bolls of the plant in red, green and blue respectively (Figure 2). The down sampling step was required as it is infeasible to apply 3d deep learning directly on high resolution LiDAR data with millions of points. Each point cloud was normalized to unit sphere to allow consistent scales among all samples in the dataset. Moreover, data augmentation was performed in each training step by rotating the input point clouds along z-axis.

Figure 1: Point cloud sample of cotton plants collected using FARO LiDAR scanner. Left point cloud represents a sample plant with original rgb values. Right point cloud represents that sample with labelled plant parts.

2.2 3D Neural architecture search

To search for network architecture with optimal segmentation performance, the 3D neural architecture search paradigm with the combination of super network training and evolutionary searching is applied as illustrated in Figure 2. The overall workflow consists of three steps. In the first step, the super network covering the candidate networks is trained. In the second step, the trained super network is used in evolutionary searching to search for top performing architecture. The searched architecture is then trained from scratch for plant part segmentation in third step.

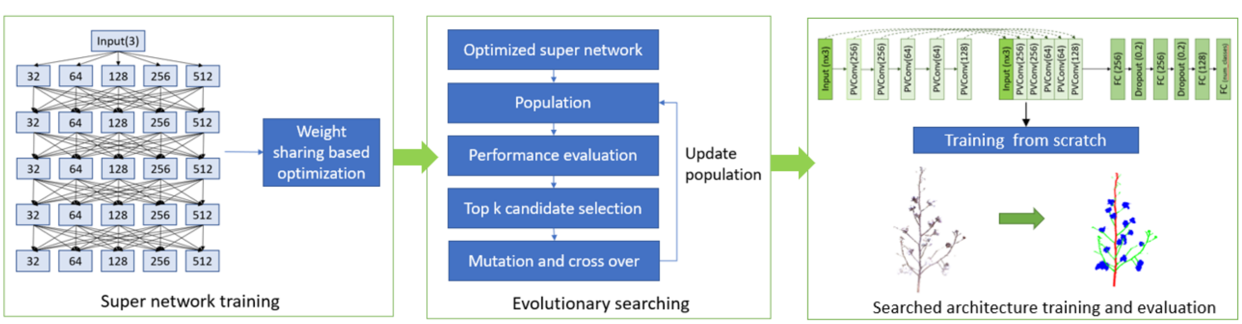

Figure 2: Workflow of neural architecture search to find optimal network for plant part segmentation.

To search for optimal network architecture for plant part segmentation, we start by defining search space covering possible candidate. We define our search space to have all networks with a maximum depth of 5. We utilize Point Voxel Convolution (PVConv) layer as the basic building block of each candidate network. While there are other feature extraction layers in 3d deep learning like set abstraction layer, edge conv layer, x conv and others. However, the PVConv provides more efficiency in terms of memory and time consumption. It utilizes point representation of point cloud for global feature extraction and voxel representation for local feature extraction. Each PVConv layer consists of two branches for point- and voxel-based input. The point branch consists of Fully connected layers while the voxel branch consists of 3d convolution layers. The extracted features from both the point and voxel branches are fused by element-wise addition operation. In each candidate network, the features from PVConv layer are extracted and passed through fully connected layers for pointwise segmentation. Search is performed for optimal number of PVConv layers in the network as well as number of output channels in each of PVConv layer in the network.

We apply 3d neural architecture search in two steps. In the first step, we formulate and optimize a super-network. The super network is designed in a way to contain all candidates as sub-networks. To perform super network training, a random candidate is selected in each training step and weights are optimized for that candidate in super network. The weight sharing is used to update weights in other parts of the super network by sharing weights of updated candidate layers with other layers at same depth in the super network.

The trained super network is performing the evolutionary searching to find network with optimal performance. We select a population of candidate networks from the super network using uniform random sampling and carry out several iterations of evolution. In each iteration, the performance of each candidate from the population is evaluated. Rather than training each candidate in the population from scratch, the weights are borrowed from the optimized super network and evaluation is performed using trained weights from super network. We select top ten high performing candidate networks and apply mutation and cross over to achieve updated the network architectures. These crossed over and mutated architectures are added to population to perform the next iteration. In the last iteration, the top performing candidate is selected as the searched candidate network. The searched network is trained from scratch for plant part segmentation. In the searched network, the features from each PVConv layer are concatenated and passed as input to the fully connected layers as per point classification.

3. RESULTS AND DISCUSSION

3.1 Searched architecture

The 3d neural architecture search applied on plant part segmentation shows that the searched architecture utilized max depth having 5 PVConv layers. In addition, the number of output channels in each PVConv layer is either 64, 128, or 256 as shown in the Figure 3.

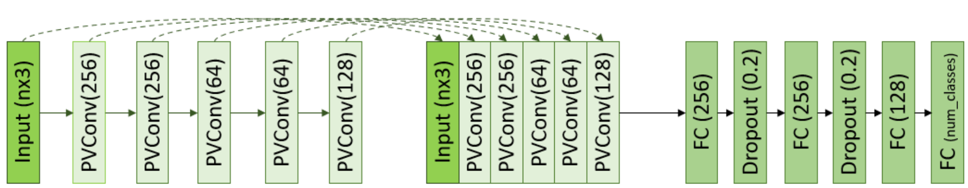

Figure 3: Searched network architecture. The number of PVConv layers as well the output channels is each PVConv layer is searched through evolutionary searched from optimized super network.

The input and features from each PVConv layer is concatenated and passed through fully connected and drop out layers to output per point scores.

3.2 Segmentation performance

The comparison with manually designed network architecture in baselines show that the searched arch outperforms in both overall and class wise segmentation performance. In terms of overall segmentation performance, the overall accuracy, mean class accuracy and mean IoU achieved by searched architecture was 2, 4 and 6 percentage points higher than the top performing baselines respectively. In terms of class-wise segmentation performance, the searched architecture outperformed baselines by more than 6 percentage point when comparing main stem and branch IoU. This margin was relatively lower (around 1%) in terms of boll IoU.

Table 1: Comparison of Mean IoU (mean iou), mean class accuracy (mAcc), overall accuracy (OA) of the searched architecture with the baselines

Network		Pointnet	Pointnet++	PVCNN	DGCNN	PointCNN	Searched architecture
Mean IoU (%)		38.61	83.19	83.88	79.88	68.79	90.04
Mean class accuracy (%)		56.51	90.52	91.24	90.09	82.9	95.07
Over all accuracy (%)		69.58	94.21	94.1	91.64	85.08	96.55
Class-wise IoU	Main stem	36.65	78.22	83.83	78.42	64.23	89.83
	Branch	9.91	75.97	73.67	70.28	60.55	83.67
	Bolls	69.29	95.38	94.13	90.94	81.57	96.61

Searched architecture made fewer mispredictions compared to baselines when visualizing inference results (Figure 4). It was observed that Pointnet++ and PVCNN mis-labelled the vertically aligned parts of branches as main stem. However the searched architecture was robust to this error. The searched architecture showed higher accuracy in boll which was mis-segmented as main stems in some parts by PointCNN and Pointnet.

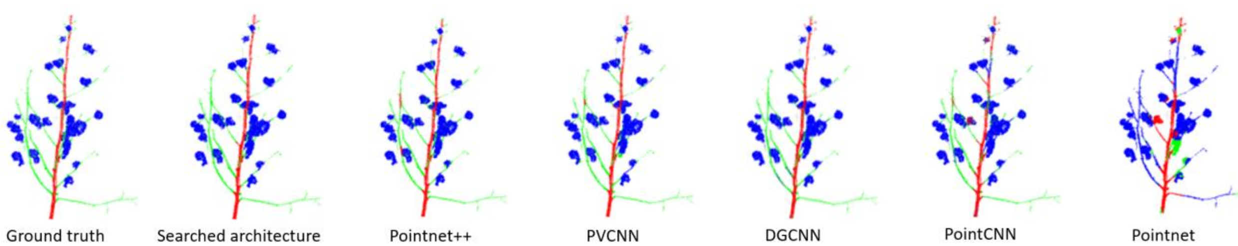

Figure 4: Visualization of inference results of searched architecture and baselines on a sample from test set and the corresponding ground truth.

4. CONCLUSIONS

Optimal deep network architecture was searched using 3d neural architecture search by combining super network training with evolutionary searching. The searched architecture outperformed the baselines with manually designed architectures in terms of both overall and class-wise segmentation performance. Overall, this method showed promising results and can be utilized to achieve optimal part segmentation of all plants including cotton. Further this can assist the plant breeder for enhancing crop yield and production and quality.

DATA AVAILABILITY STATEMENT

The collected dataset used in this study is available at https://doi.org/10.25739/vnr9-xt59.

ACKNOWLEDGMENTS

Authors gratefully thank Dr. Shangpeng Sun and Javier Rodriguez for data collection. Authors additionally thank Bio-sensing and Instrumentation Lab (BSAIL) members for their helpful discussions. Authors further gratefully thank for computing resources and technical expertise from Georgia Advanced Computing Resource Center (GACRC) and Reasoning and Knowledge Discovery (RISE) Lab, University of Georgia. This work was partially supported by the Cotton Incorporated, National Science Foundation (grant number: 1934481) and Georgia Cotton Commission. The first author is financially supported by Higher Education Commission under US-Pakistan Knowledge Corridor.

REFERENCES

[1] Xiang, Lirong, Yin Bao, Lie Tang, Diego Ortiz, and Maria G. Salas-Fernandez. "Automated morphological traits extraction for sorghum plants via 3D point cloud data analysis." Computers and Electronics in Agriculture 162 (2019)

[2] Miao, Teng, Chao Zhu, Tongyu Xu, Tao Yang, Na Li, Yuncheng Zhou, and Hanbing Deng. "Automatic stem-leaf segmentation of maize shoots using three-dimensional point cloud." Computers and Electronics in Agriculture 187 (2021)

[3] Bao, Yin, Lie Tang, Srikant Srinivasan, and Patrick S. Schnable. "Field-based architectural traits characterisation of maize plant using time-of-flight 3D imaging." Biosystems Engineering 178 (2019)

[4] Elnashef, Bashar, Sagi Filin, and Ran Nisim Lati. "Tensor-based classification and segmentation of three-dimensional point clouds for organ-level plant phenotyping and growth analysis." Computers and electronics in agriculture 156 (2019)