SEMANTIC SEGMENTATION OF UAV LIDAR DATA FOR TREE PLANTATIONS

Jinyuan Shao[1], Ayman Habib[2], Songlin Fei[1] *

[1] Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47907, USA - (jyshao, sfei)@purdue.edu
[2] Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, 47907, USA - (ahabib)@purdue.edu

KEY WORDS: LiDAR, UAV, Forestry, Point Cloud, Deep Learning.

ABSTRACT:

Tree plantations, characterized by large-scale cultivation of trees with high commercial values, often rely on accurate inventory data to improve their capacity. However, understanding tree plantations with different components on a large scale for growth prediction is still a tricky problem. In this paper, we harness the power of Unmanned Aerial Vehicle (UAV) Light Detection and Ranging (LiDAR) systems to acquire 3D point clouds of tree plantations and investigate the potential of deep learning segmentation for enhanced understanding of plantation UAV LiDAR point clouds, thereby promoting precision forest management. Two datasets from the same plantation without debris on the ground and with harvested debris were tested. Experimental results showed that we were able to process a plantation consisting of 300 trees in 2 min and achieve an overall accuracy of 95% segmentation for this plantation. This research demonstrates the feasibility of the deep learning method in segmenting large-scale tree plantation point clouds, which is able to speed up the inventory of tree plantations.

1. INTRODUCTION

A significant amount of forest products are derived from different species of trees that are cultivated from tree plantations (Shmulsky and Jones, 2019). To accurately evaluate the economic values of tree plantations, tree inventory that provides individual trees' timber volume is essential (Klimas et al., 2012). Traditional forest inventory methods, including manual tree counting and diameter measurement at breast height (DBH), are time-consuming, labor-intensive, and subject to human errors.

In recent years, advancements in remote sensing technologies have provided promising alternatives for forest inventory (White et al., 2016). Among these, UAV LiDAR systems have emerged as a powerful tool for collecting precise three-dimensional (3D) data over extensive areas in a relatively short period of time (Wallace et al., 2012). By integrating with UAV, LiDAR is able to measure distances, penetrate the forest canopy in the air, and provide highly accurate data about the ground surface and the structures on it (Lin et al., 2022). As a result, it holds great potential for improving the efficiency and accuracy of forest inventory tasks (Wallace et al., 2012).

To conduct tree plantation inventory with LiDAR, it's essential to derive different components of tree plantations accurately, which relies on LiDAR point cloud segmentation. Conventional point cloud segmentation requires hand-crafted geometric feature extraction, such as linear, planar, and scatter features. And then these fundamental geometric features are further clustered into meaningful objects, such as ground, trunks, and canopy. Geometric features were analyzed and trunks were extracted by using cylinder fitting for pine, spruce, and birch, and 73% stem mapping accuracy was achieved (Liang et al., 2011). Other geometric features were also able to be used for stem points segmentation. By computing surface curvature for each point, tree stems were further extracted with postprocessing for a pure Chinese scholar (*Styphnolobium japonicum*) tree plantation (Zhang et al., 2019). To avoid point-wise geometric features calculation, a bottom-up approach that utilized a peak detection algorithm for stem localization in height normalization point clouds was developed, and the experimental results showed that this approach performed better than DBSCAN (Ester et al., 1996) for red oaks stem localization (Lin et al., 2021).

Segmented stem points can be used to localize individual stems, but it is unable to derive other inventory attributes. With more structural forest attributes being requested, segmenting only stem points is apparently not sufficient. Therefore, algorithms that are able to understand point clouds semantically were subsequently adopted in forestry. A graph-based method that relied on all points' geometric features classified tropical trees into two classes (i.e., leaf and wood), and then woody components were reconstructed to obtain quantitative structural models (QSM) and estimated biomass for tropical trees (Wang et al., 2020). These geometric feature-based methods require point-wise computation, which leads to high time complexity. Moreover, it is difficult for geometric feature-based methods to capture global context information and understand semantic information.

As deep learning showed its capability in point cloud understanding (Qi et al., 2017), some studies explored the potential of deep learning for forest point cloud segmentation. A PointNet model (Qi et al., 2017) was adopted to classify forest point clouds into four categories and proved deep learning-based method is sensor agnostic. Raw measurements provided by laser scanners were also used as the input of a deep neural net for boreal forest segmentation (Kaijaluoto et al., 2022). Although deep learning boosted semantic segmentation for forest LiDAR point cloud, all of the existing research was developed based on algorithms that were initially developed for a single object or on a small scale, which is not suitable for the outdoor forest environment. To solve this problem, large-scale LiDAR point clouds were tiled into small blocks (i.e., 1 x 1 m). However, this strategy destroys the spatial relationship and it's hard to capture the global features of forest point clouds. Some research optimized algorithms for this problem (Hu et al., 2020), but the effectiveness has not been verified in forest environ-

* Corresponding author

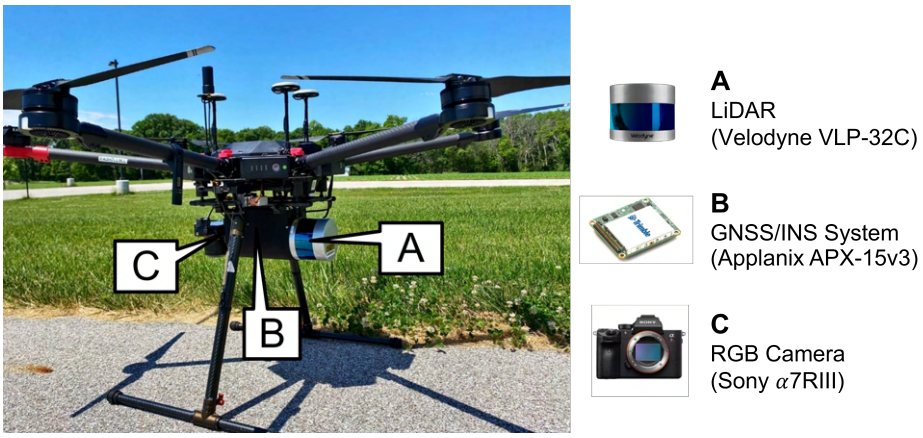

Figure 1. UAV LiDAR system for plantation data acquisition.

ments.

In this paper, we apply a superpoint-based deep learning method to UAV forest point clouds to overcome these challenges. UAV LiDAR data collected from large-area tree plantations with a massive number of points is the input. Two datasets from a northern red oak (*Quercus rubra*) plantation with and without debris were tested. This research contributes to ongoing efforts to improve precision forest management by demonstrating a practical approach to enhance the speed and accuracy of forest inventory data acquisition.

2. METHODOLOGY

2.1 UAV LiDAR System

We utilized an in-house UAV LiDAR system, as illustrated in Figure 1. This UAV LiDAR system consists of a multi-beam rotating laser scanner, an RGB camera, and a GNSS/INS module for direct georeferencing. The VLP-32C laser scanner at the system's core employs 32 radially arranged laser rangefinders. Its vertical and horizontal field of view (FOV), as observed from the LiDAR unit's rotational axis, spans 40° (extending from $+15^\circ$ to -25°) and 360°, respectively. This scanner collects approximately 600K points per second (under single return mode), maintaining a range precision of $\pm$ 3 cm and a maximum reach of 200 m. The GNSS/INS module is expected to deliver post-processing positional precision within a range of $\pm$ 2 to $\pm$ 5 cm, and the accuracy for roll/pitch and heading is estimated to be $\pm$ 0.025° and $\pm$ 0.08° respectively.

The UAV system is engineered such that the rotational axis of the LiDAR unit remains nearly parallel to the direction of flight. The FOV across the flight path was established to be $\pm 70^\circ$ from the nadir, meaning a point is only reconstructed if the laser beam's pointing direction falls within $\pm 70^\circ$ of the nadir. Utilizing a rotating multi-beam LiDAR unit provides several advantages, including its distinctive scanning mechanism. Given the laser beams' multidirectional rotation and firing, the likelihood of LiDAR energy infiltrating foliage and mapping features beneath the canopy is considerably heightened. Additionally, this method helps alleviate occlusion issues as multiple laser beams can capture a single location in the object space at different times. To capitalize on this unique scanning technique and form point clouds with an extensive swath along the flight direction, stringent system calibration is required. Therefore, this study employed an in-situ system calibration method to ascertain the relative orientation and position between the GNSS/INS unit and the onboard sensors (Ravi et al., 2018).

2.2 Study Site and Dataset

The study site, Plot 5A, is a plantation of the northern red oak species (*Quercus rubra*) situated within the bounds of Martell forest (Figure 2). Plot 5A is a research-oriented forest in the state of Indiana, USA. The plantation in Plot 5A was planted in 2007 following a carefully planned grid pattern. This pattern consists of 22 rows each spaced apart by 5 meters, and every row accommodating 50 trees distanced at 2.5 meters from each other. In order to obtain robust and detailed data, our in-house UAV LiDAR was employed. One notable differential factor between the two sets of data is the varying tree density. Following the first data acquisition, a portion of trees within Plot 5A was thinned, hence causing a decrease in tree numbers prior to the second data acquisition. At the point of the second scanning, debris of dead trees could still be found scattered on the forest floor.

To ensure a sufficient amount of data for both training and evaluation, a dataset was generated from the data collected in Plot 5A. Each LiDAR point within this dataset was carefully annotated with corresponding semantic labels. This is in response to the complex nature of Plot 5A and also in preparation for conducting a variety of objective research analyses in the future. We defined five distinctive semantic categories that each represent different functional components within forests. These include the categories of ground, debris, trunks, crowns, and others. Here are the definitions and associated functions for each category:

- **Ground**: Forest floor in the LiDAR point clouds. The segmentation result of the ground provides the basic information for the generation of Digital Terrain Models (DTMs).
- **Debris**: Debris includes dead materials that lie on the ground, which could serve as potential fuels for forest fires or wildlife habitats.
- **Trunks**: Trunks are parts of the trees that locate below the Crown Base Height (CBH). The CBH is defined as the distance from the ground surface to the lowest live branch within a tree crown (Popescu and Zhao, 2008). Trunks are the main sources of timber, fiber, and carbon.

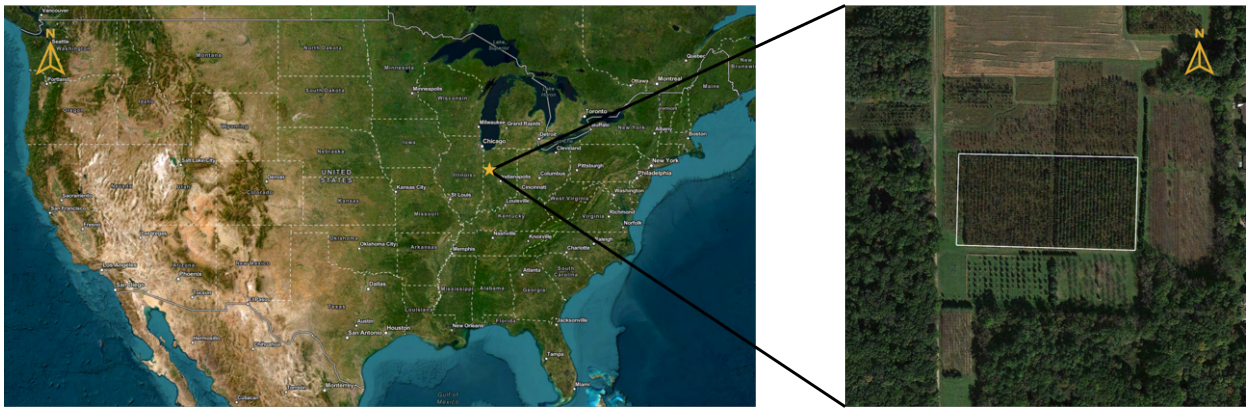

Figure 2. The study site, a northern red oak plantation, is located in West Lafayette, Indiana, USA.

- **Crowns**: Crowns are the canopies of the trees that are located above the CBH.
- **Others**: This category encompasses unrecognized points in the forests, such as noise and shrubs.

2.3 Semantic Segmentation of Point Cloud

To allow for semantic segmentation with massive LiDAR points input, Superpoint Graph (SPG) (Landrieu and Simonovsky, 2018), which is a deep learning method designed for large-scale point cloud segmentation is adopted. The process consists of two main stages: partitioning and learning. In the partitioning stage, the input point cloud is initially clustered into a set of small, geometrically-homogeneous primitives denoted here forth as superpoints. A graph is then constructed by connecting neighboring superpoints with superedges, resulting in a superpoint graph. In the learning stage, a PointNet (Qi et al., 2017) is employed to extract superpoint embedding from the superpoints. Finally, a graph-based neural network is then used to process superpoints embeddings and superedges and output semantic labels for each superpoint.

In the geometric partitioning (Figure 3), geometric features are first calculated, including linearity, planarity, and scattering. For each point, eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are computed within its neighborhood. Then the geometric classification can be derived according to the largest of the following (Demantké et al., 2012):

$$\begin{aligned} \textit{Linearity} &= \frac{\lambda_1 - \lambda_2}{\lambda_1} \\ \textit{Planarity} &= \frac{\lambda_2 - \lambda_3}{\lambda_1} \\ \textit{Scattering} &= \frac{\lambda_3}{\lambda_1} \end{aligned} \tag{1}$$

Other than the above three features, elevation is also included as one of the geometric features. Then, a graph structure is constructed for the entire point cloud. The geometric partitioning can be performed based on this graph and it is defined as the following optimization problem:

$$\arg\min_{g \in R^{dg}} \sum_{i \in C} \|g_i - f_i\|^2 + \mu \sum_{(i,j) \in E_{\text{nn}}} w_{i,j} \left[g_i - g_j \neq 0 \right] \tag{2}$$

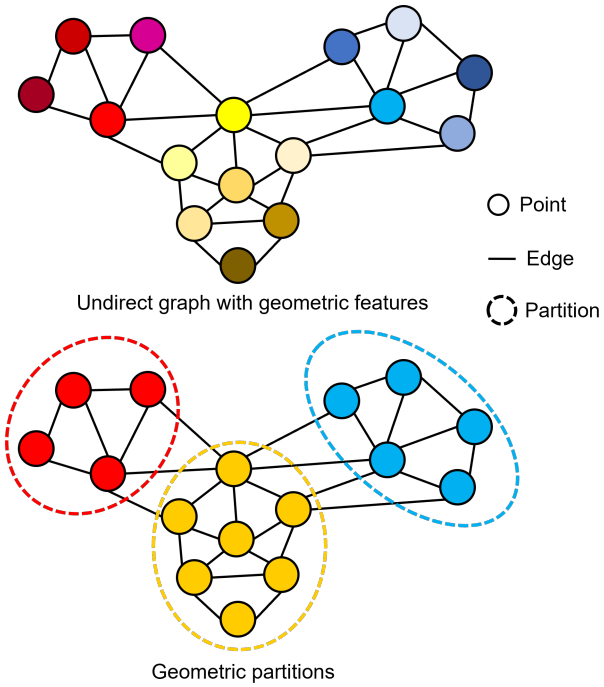

Figure 3. Illustration of geometric partitioning. Top: an undirect graph of point cloud, nodes are colored by geometric features. Bottom: partitions that are piecewise constant approximations of the geometric features.

where i and j are points in the LiDAR point cloud; g is a piecewise constant approximation of the geometric feature f; ω is the edge weight that is related to the distance between points; μ is the regularization strength. The first part of summation represents the fidelity that makes components of g correspond to homogeneous values of f; the second part adds a penalty term for each introduced geometric partitioning. The regularization strength μ determines the trade-off between fidelity and simplicity and thus simplicity determines the number of clusters. Each cluster in the partition corresponds to a superpoint, and their adjacency edges are referred to as superedges.

The learning stage aims to classify every superpoints. Although the construction of SPG reduces the complexity of computation for deep learning input, the sizes of superpoints are different so they need to be converted to a unified form. Therefore, a PointNet (Qi et al., 2017) is introduced to extract superpoint

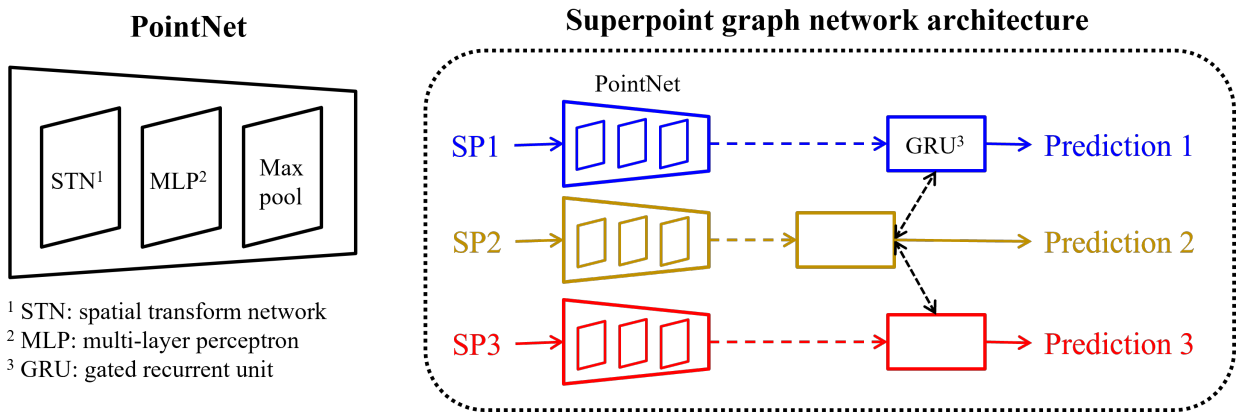

Figure 4. Model structure of superpoint graph. Each suerpoint is fed into a PointNet to get embedding vector, then superpoint embeddings and superedges are fed into a graph net that consists of a series of GRU.

embedding that provides a rich representation of the underlying 3D structure. Figure **??** illustrates the structure of PoineNet. A single partition of point clouds is first fed into a T-Net, a module to spatial transform network which is to rotate and translate point clouds to a fixed position that is spatial invariant. Then the point cloud data is processed by a series of Multi-Layer Perceptrons (MLPs) and max-pooling operations to get an embedding, a descriptor for the superpoint.

A graph convolutional network is used to learn SPG (Simonovsky and Komodakis, 2017). As demonstrated in Figure 4, each superpoint is embedded by a Pointnet (Qi et al., 2017), and each superedge is encoded by the offset and the ratio of shape and size between adjacent superpoints. Superpoint embeddings are then fed into a Gated Recurrent Unit (GRU) by message passing with superedges. Superedge's weight is from the output of a multi-layer perceptron. A GRU leverages two gate units, "reset" and "update", to discern how much past information should be kept or forgotten (Cho et al., 2014). This capability of selective memory, forgetting irrelevant data while concentrating on informative features, enables a more compact representation. Finally, the output of SPG'node is the segmentation result of point clouds.

3. RESULTS

To assess the performance of the SPG model, we selected two testing sets: a tile of 60 x 70 m from the 2021 data and a tile of 65 x 50 m from the 2022 data. The testing set of 2021 has 300 trees with 25 columns and 12 rows, while the testing set of 2022 has 180 trees after harvesting. The remained data was utilized for training and validation. Evaluation metrics included commission error (precision), omission error (recall), F1-score, and Overall Accuracy (OA), as well as overall execution time. Table 1 and Table 2 present the quantitative results for the 2021 and 2022 datasets, respectively.

The results indicate that the SPG model performs well in differentiating ground, trunks, and crowns for both datasets. For the 2021 data without debris (Table 1), the model achieved perfect precision in detecting ground and crowns, with F1-scores of 1.00 and 0.91, respectively. In the case of trunks, the model demonstrated high recall (0.96) but lower precision (0.41), resulting in an F1-score of 0.58. The overall accuracy for the 2021 data was 95%. The model did not identify any debris or other features, which is consistent with the absence of debris in this dataset. For the 2022 data with debris (Table 2), the SPG model maintained high performance in detecting ground, trunks, and crowns, with F1-scores of 0.98, 0.69, and 0.97, respectively. Notably, the model successfully detected debris with a precision of 0.82, recall of 0.72, and an F1-score of 0.77. The overall accuracy for the 2022 dataset also stood at 95%. Similar to the 2021 dataset, the model did not identify any other features in the 2022 data. In terms of execution efficiency, SPG consumed 115 s for 2021 dataset whose area is 4200 m^2, and 98 s for the 2022 dataset with 3250 m^2. Figure 5 and Figure 6 show the visualization of segmentation results for 2021 and 2022 datasets, respectively.

Table 1. Quantitative segmentation results of tree plantation without debris.

	Precision	Recall	F1-score
Ground	1.00	1.00	1.00
Debris	0.00	N/A	N/A
Trunks	0.41	0.96	0.58
Crowns	1.0	0.84	0.91
Others	0.00	N/A	N/A
Running time	115 seconds		
Overall accuracy	95%		

Table 2. Quantitative segmentation results of tree plantation with debris.

	Precision	Recall	F1-score
Ground	0.97	0.98	0.98
Debris	0.82	0.72	0.77
Trunks	0.55	0.91	0.69
Crowns	1.00	0.94	0.97
Others	0.00	N/A	N/A
Running time	98 seconds		
Overall accuracy	95%		

4. CONCLUSION

Precision management of tree plantations requires an accurate tree inventory. In this study, we implemented a deep learning-based model, SPG, for the semantic segmentation of point clouds covering tree plantations. By leveraging our in-house UAV LiDAR system, we collected point cloud data for the entire red oak plantation. The segmentation results on the testing

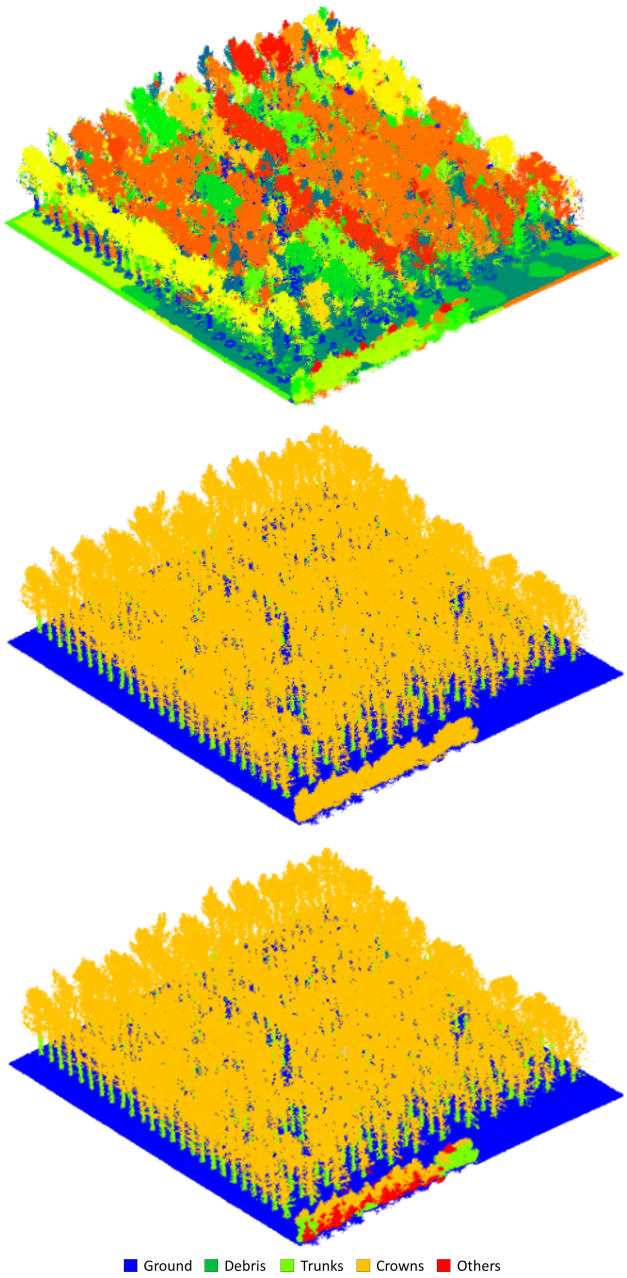

Figure 5. Visualization of segmentation result for 2021 UAV data without debris. From top to bottom: geometric partitioning, reference, and prediction.

set showed that SPG is able to segment a plantation that consists of 300 trees and achieve 95% overall accuracy in 2 min. The SPG is not only able to recognize trunk and crown points but also able to detect harvested trees on the ground. Our experimental results proved the promising potential of UAV LiDAR systems for tree plantation inventory and the efficiency of deep learning for plantation point cloud segmentation. In the future, we'll expand our study to complex natural forests and explore some potential downstream applications, such as stem mapping and diameter measuring.

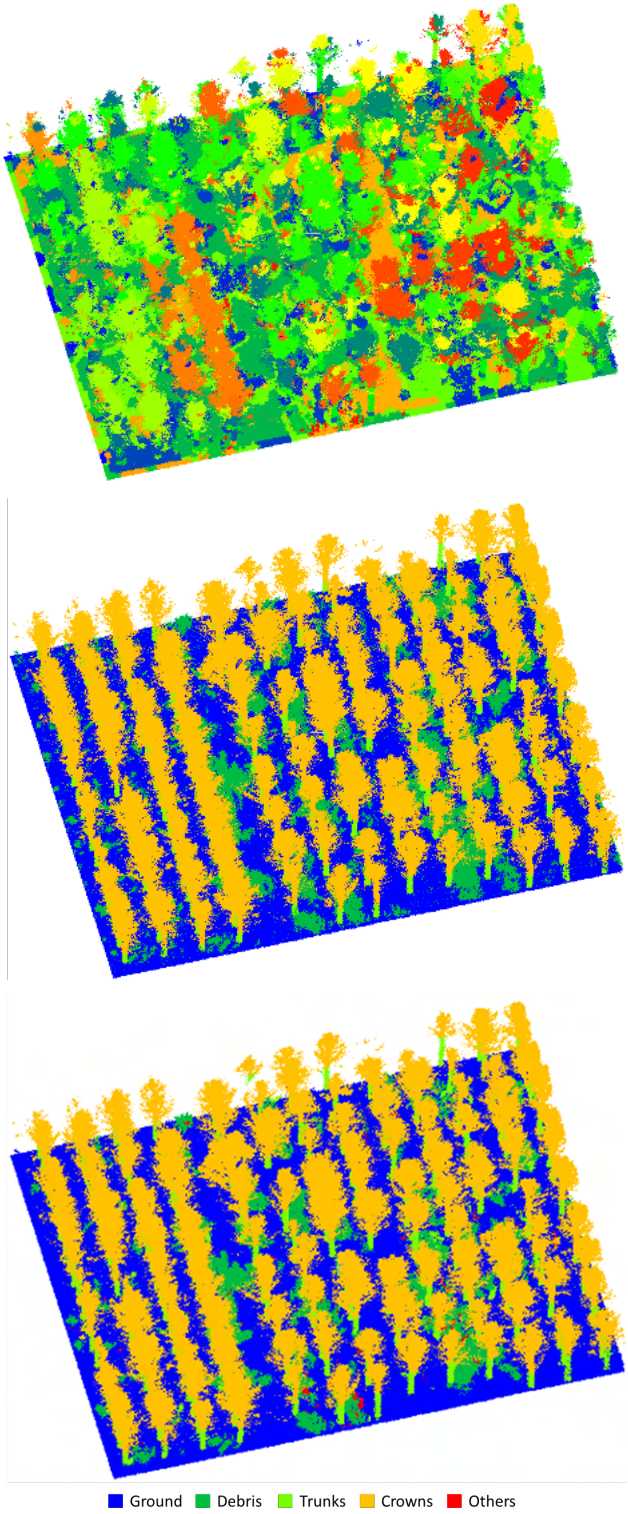

Figure 6. Visualization of segmentation result for 2021 UAV data with debris.From top to bottom: geometric partitioning, reference, and prediction.

5. ACKNOWLEDGEMENT

Funding for this research is partially supported by USDA NIFA # 2023-68012-38992.

REFERENCES

Cho, K., Van Merriënboer, B., Bahdanau, D., Bengio, Y., 2014. On the properties of neural machine translation: Encoder-decoder approaches. *arXiv preprint arXiv:1409.1259*.

Demantké, J., Mallet, C., David, N., Vallet, B., 2012. Dimensionality based scale selection in 3D lidar point clouds. *The international archives of the photogrammetry, remote sensing and spatial information sciences*, 38, 97–102.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X. et al., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. *kdd*, 96number 34, 226–231.

Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., Trigoni, N., Markham, A., 2020. Randla-net: Efficient semantic segmentation of large-scale point clouds. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 11108–11117.

Kaijaluoto, R., Kukko, A., El Issaoui, A., Hyyppä, J., Kaartinen, H., 2022. Semantic segmentation of point cloud data using raw laser scanner measurements and deep neural networks. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, 3, 100011.

Klimas, C. A., Kainer, K. A., de Oliveira Wadt, L. H., 2012. The economic value of sustainable seed and timber harvests of multi-use species: an example using Carapa guianensis. *Forest Ecology and Management*, 268, 81–91.

Landrieu, L., Simonovsky, M., 2018. Large-scale point cloud semantic segmentation with superpoint graphs. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Liang, X., Litkey, P., Hyyppa, J., Kaartinen, H., Vastaranta, M., Holopainen, M., 2011. Automatic stem mapping using single-scan terrestrial laser scanning. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2), 661–670.

Lin, Y.-C., Liu, J., Fei, S., Habib, A., 2021. Leaf-off and leaf-on uav lidar surveys for single-tree inventory in forest plantations. *Drones*, 5(4), 115.

Lin, Y.-C., Shao, J., Shin, S.-Y., Saka, Z., Joseph, M., Manish, R., Fei, S., Habib, A., 2022. Comparative Analysis of Multi-Platform, Multi-Resolution, Multi-Temporal LiDAR Data for Forest Inventory. *Remote Sensing*, 14(3), 649.

Popescu, S. C., Zhao, K., 2008. A voxel-based lidar method for estimating crown base height for deciduous and pine trees. *Remote sensing of environment*, 112(3), 767–781.

Qi, C. R., Su, H., Mo, K., Guibas, L. J., 2017. Pointnet: Deep learning on point sets for 3d classification and segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Ravi, R., Lin, Y.-J., Elbahnasawy, M., Shamseldin, T., Habib, A., 2018. Simultaneous system calibration of a multi-lidar multicamera mobile mapping platform. *IEEE Journal of selected topics in applied earth observations and remote sensing*, 11(5), 1694–1714.

Shmulsky, R., Jones, P. D., 2019. Forest products and wood science: an introduction.

Simonovsky, M., Komodakis, N., 2017. Dynamic edge-conditioned filters in convolutional neural networks on graphs. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3693–3702.

Wallace, L., Lucieer, A., Watson, C., Turner, D., 2012. Development of a UAV-LiDAR system with application to forest inventory. *Remote sensing*, 4(6), 1519–1543.

Wang, D., Momo Takoudjou, S., Casella, E., 2020. LeWoS: A universal leaf-wood classification method to facilitate the 3D modelling of large tropical trees using terrestrial LiDAR. *Methods in Ecology and Evolution*, 11(3), 376–389.

White, J. C., Coops, N. C., Wulder, M. A., Vastaranta, M., Hilker, T., Tompalski, P., 2016. Remote sensing technologies for enhancing forest inventories: A review. *Canadian Journal of Remote Sensing*, 42(5), 619–641.

Zhang, W., Wan, P., Wang, T., Cai, S., Chen, Y., Jin, X., Yan, G., 2019. A novel approach for the detection of standing tree stems from plot-level terrestrial laser scanning data. *Remote sensing*, 11(2), 211.