Automated Plant Species Identification Using 3D LiDAR Point Clouds: A Case Study Using Cabbage and Maize Plants

Mukesh Kumar Verma[1,], Manohar Yadav[1]*

[1] Geographic Information System (GIS) Cell, Motilal Nehru National Institute of Technology, Prayagraj, India, mukeshverma02feb@gmail.com, ssmyadav@mnnit.ac.in@gmail.com

* corresponding author

doi: 10.5281/zenodo.11555370

Abstract: Automated plant species identification plays a crucial role in various ecological, agricultural, and environmental monitoring applications. This paper proposes an effective approach utilizing 3D LiDAR point clouds for automated plant species identification. LiDAR technology provides detailed and accurate spatial information about the vegetation canopy, enabling precise characterization of plant structures in three dimensions. Our methodology involves several key steps. Initially, raw 3D LiDAR point cloud data is acquired from the two study areas (i.e., maize crop field and cabbage field). Next, preprocessing techniques are applied to filter noise and extract relevant features from the point cloud. Finally, plant volume, projected leaf area and plant heights of the two agricultural crops are estimated to differentiate the plant species on the basis of structural information. One of the significant advantages of our proposed approach is its ability to capture the intricate structural characteristics of different plant species with high accuracy and efficiency. By leveraging 3D LiDAR technology, our method transcends the limitations of traditional 2D imaging techniques, which often struggle to accurately differentiate between species with similar visual appearances. Experimental results demonstrate the effectiveness of our approach in accurately identifying plant species from 3D LiDAR point clouds. The proposed method shows promising performance across two types of vegetation and can further extended to other types of crops within the same field.

Keywords: 3D point cloud; plant phenotyping; terrestrial laser scanning; total station.

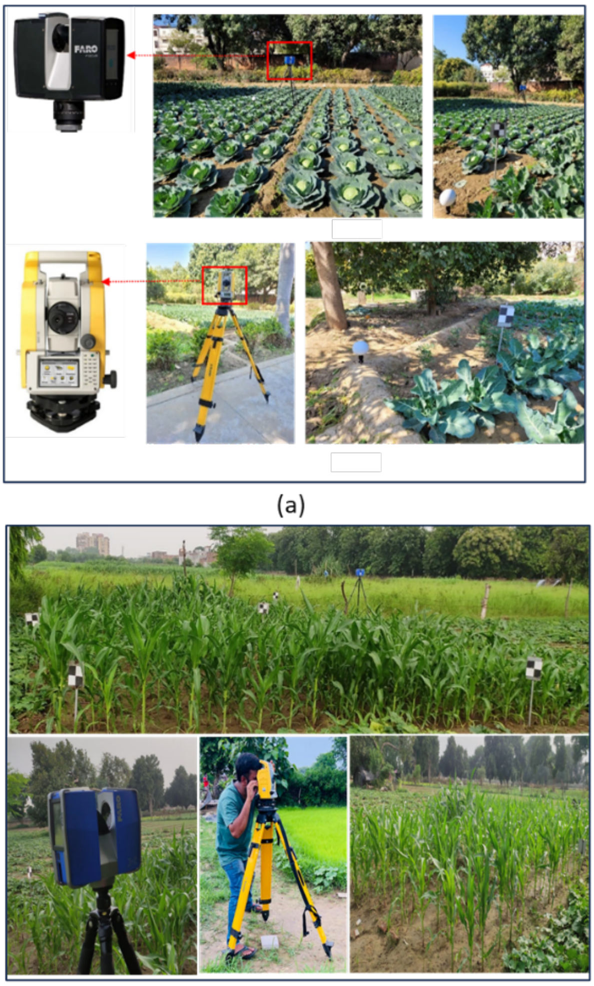
1 Introduction

Plant phenotyping is essential for plant species identification (Verma and Yadav 2024). Automated plant species identification using LiDAR technology is a promising approach for precise forest management and phenotyping. Various studies have demonstrated the effectiveness of LiDAR data in classifying tree species based on features like roughness parameters (Ana et al. 2022), scan angles (Brindusa et al. 2022), deep learning models (Bingjie et al. 2022) and intensity and texture features (Ao et al. 2022). These methods utilize advanced techniques such as K-means clustering, convolutional neural networks (CNNs), and random forest (RF) algorithms to accurately classify different plant species. LiDAR data enables the extraction of 3D structural information, individual tree point clouds, and phenotypic traits like leaf area and stem position, contributing to high-throughput phenotyping and precision agriculture applications (Lombard et al. 2020). By combining these diverse approaches, automated plant species identification using LiDAR proves to be a valuable tool for enhancing forest inventory, species discrimination, and plant phenotyping. We have developed methodology for the plant species identification using plant phenotyping parameters such as plant height, plant volume and projected leaf area using terrestrial laser scanner (TLS) and total station (TS) instruments.

2 Materials and methods

2.1 Study areas

The study area 1 chosen for the experiment and performance assessment of proposed methodology was located in Narayani Ashram (25° 29' 49.16273" N, 81° 52' 6.52379" E), a place in Govindpur area, Prayagraj, India.



(a)

(b)

Figure 1. Data collection from study areas (a) study area 1 (b) study area 2.

The selected study area 1 was a cabbage field with sandy loam soil that covered an area of 450 m2 (30 m × 15 m). Another study area 2 was chosen near Sam Higginbottom University of Agriculture, Technology and Sciences (25° 24' 29.9945" N, 81o 49' 56.6233"), Prayagraj, India. This study area 2 was an agricultural field having maize crop. Data collection was performed using TLS and TS. The maize crop in this study area was analyzed throughput its crop period and data collection were done at the tasseling stage of the maize crop. Figure 1 presents the study area 1 and 2.

2.2 Data acquisition

The data at two study areas were acquired using the experimental setup provided in the Figure 2. Five scan stations were setup to collect the complete 3D structure of the crops using FARO Focus^{3D} Laser Scanner. The registration targets such as Spheres and Checkerboards are placed within the experimental setup to facilitates the registration process. The center of the checkerboards was also measured using the Trimble M3 Total Station. As center of checkerboards were measured using both the above instruments. Coordinate transformation is mandatory as we need the alignment of the z-axis towards the plumb line for the estimation of morphological parameters of the vegetable crops.

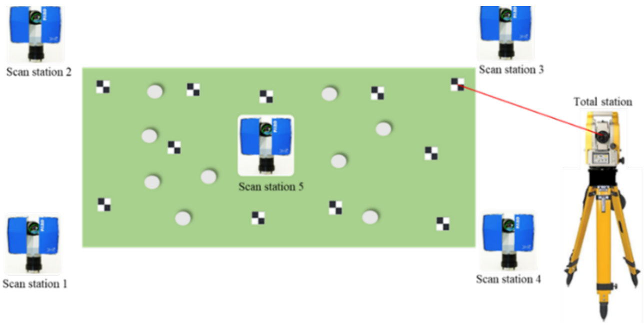


Figure 2. Experimental setup for data collection from study area 1 and 2.

The ground truth data was collected from the cabbage and maize crop field using measuring tape. The plant height, plant width and plant length were measured manually. And for maize crops the same parameters were measured. Projected leaf area and plant volume were calculated using the formulas given as equation (1) and (2) (Verma and Yadav 2024):

$$\text{Projected leaf area} = \pi * \frac{l}{2} * \frac{w}{2} \tag{1}$$

$$\text{Plant volume} = \frac{3}{4} * \pi * \frac{l}{2} * \frac{w}{2} * \frac{h}{2} \tag{2}$$

2.3 Proposed methodology

The data was collected from the study area 1 and 2 using instruments FARO Focus3D Laser Scanner and Trimble M3 Total Station. After the data collection based on the experimental setup in Figure 2, the raw point cloud data was merged and transformed using the FARO Scene software. Then, ground points were filtered using the methodology provided by Yadav et al. (2021). Then k-means algorithm was implemented to segment each individual from the non-ground points. Figure 3 presents the methodology workflow for plant species identification.

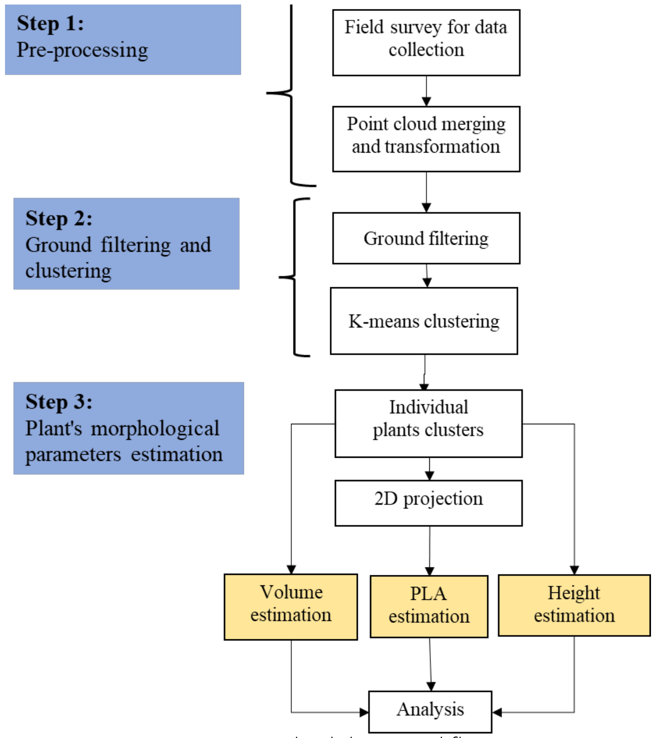


Figure 3. Methodology workflow.

2.3.1 *Plant height estimation*

The height of each plant was determined by averaging the distances between ten points located at both the top and bottom extremities of the plant. These averages were then subtracted to obtain the final height measurement, ensuring robustness against potential irregularities in plant morphology. Mathematically, the height (H) of each plant was calculated using the equation (3):

$$H = \frac{1}{10}\left(\sum_{i=1}^{10} P_i^{top}\right) - \frac{1}{10}\left(\sum_{i=1}^{10} P_i^{bottom}\right) \tag{3}$$

Where P_i^{top} and P_i^{bottom} represents the i-th point from the top and bottom of the plant, respectively. This meticulous approach ensured accurate and reliable height estimations for each individual plant, laying a solid foundation for subsequent analyses in our research.

2.3.2 *Projected Leaf Area and Volume*

The Quickhull algorithm offers an efficient method for computing the convex hull of a set of points in a plane. Initially, the algorithm identifies the leftmost (P_{left}) and rightmost (P_{right}) points as initial endpoints. These endpoints are then used to form a triangle encompassing all other points in the set. The process involves strategically partitioning the point set into subsets based on their position relative to the triangles formed by these initial endpoints and additional points. This partitioning is achieved by calculating the distance of each point (Pi) from the line segment connecting Pleft and Pright, represents the equation (4):

$$Hd_i = \frac{|(P_{right} - P_{left}) * (P_i - P_{left})|}{||P_{right} - P_{left}||} \tag{4}$$

Where × denotes the cross product and ||.|| denotes the Euclidean norm. Points with the maximum distance (dmax) are identified as the ones lying outside the current triangle. The process continues recursively until no points lie

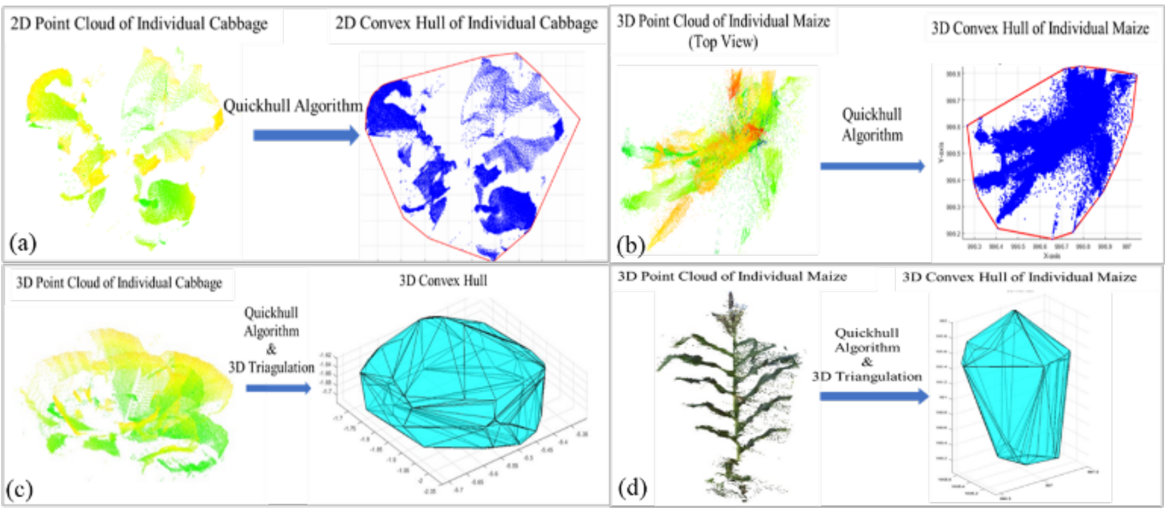


Figure 4. 2D convex hull of the plants after projection.

outside the current triangle. The process continues recursively until no points lie outside the triangles, at which point the convex hull is formed by merging the hulls of the subsets. This can be mathematically represented as equation (5):

$$\text{ConvexHull} = \text{Merge}(\text{QuickHull}(P_{left}, P_{right}, P_{outside})) \quad (5)$$

Where:

P_{left}: Points on the left side of a dividing line or plane.

P_{right}: Points on the right side of the same dividing line or plane.

$P_{outside}$: Points that lie outside the current convex hull being considered.

This iterative approach efficiently constructs the convex hull, making the Quickhull algorithm a valuable tool in computational geometry. Figure 4 presents the 2D convex hull of the crops.

3 Results

The plotted line graph (Figure 5) depicting cabbage plant height, projected leaf area, and volume, utilizing 3D point cloud data. At maturity, parameters were estimated to offer insights into plant development. These findings emphasize the complexity of cabbage plant growth and make evident the value of 3D point cloud analysis in understanding plant development dynamics. Figure 5 presents the variation in plant height, projected leaf area and volumes of the plants.

The stressed plants from the above plots (Figure 5) can be detected if the height of the plant is large and volume is less or large area is giving the less volume comparatively. A healthy plant should have optimized values of height, area and volumes. By comparing the respective height, area and volume parameters the plant species can be identified automatically. The variation in the plant height, volume and area can be easily differentiated among the cabbage and maize crops to differentiate between cabbage and maize crops. This study is a case study using cabbage and maize crops, further this methodology can be extended to differentiate between the greater number of crops automatically.

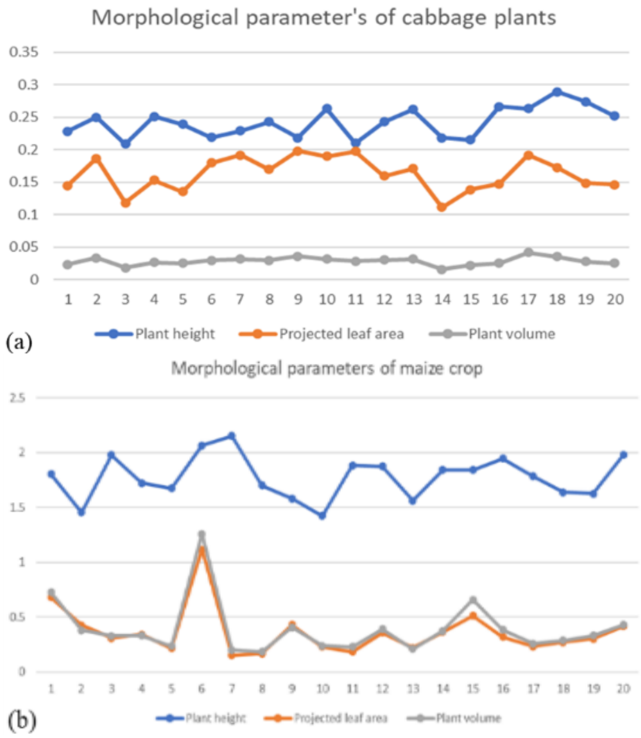


Figure 5. (a) Variation in the Plant Height, Projected Leaf Area and Volume of the individual cabbage plants (20 Sample Plants), (b) Variation in the Plant Height, Projected Leaf Area and Volume of the individual maize plants (20 Sample Plants).

4 Discussion

The study aimed to assess TLS data for mapping cabbage (low vegetation) and maize (high vegetation) by evaluating plant height, leaf area, and plant volume. TLS provided accurate estimations for cabbage but struggled with maize's complex foliage structure, impacting accuracy. Resolution influenced TLS's suitability, being advantageous for cabbage but insufficient for maize. Penetration depth into dense foliage, processing complexity, cost, and repeatability were significant considerations. Despite TLS's potential, challenges like resolution and processing complexity need addressing for effective monitoring. Future research should focus on improving TLS resolution, processing efficiency, and integrating it with other remote sensing technologies for comprehensive agricultural vegetation monitoring.

5 Conclusions

This study explores the potential of TLS for mapping vegetation in agricultural landscapes, focusing on cabbage and maize. TLS accurately estimates vegetation parameters in low vegetation like cabbage but faces challenges with high vegetation like maize due to resolution limitations. Despite hurdles, TLS offers advantages like rapid, non-destructive data acquisition, enhancing vegetation monitoring in agriculture. Integrating TLS with other remote sensing tech like LiDAR and multispectral imaging shows promise for comprehensive vegetation mapping. Addressing TLS limitations requires ongoing research and technology improvements. Overall, TLS is a valuable tool for agricultural vegetation monitoring, with continued innovation vital for addressing agriculture challenges like crop monitoring and ecosystem management.

6 References

Ao, Z., Wu, F., Hu, S., Sun, Y., Su, Y., Guo, Q. Xin, Q., 2022. Automatic segmentation of stem and leaf components and individual maize plants in field terrestrial LiDAR data using convolutional neural networks. The Crop Journal 10 (5), 1239-1250.

Budei, B.C., St-Onge, B., Fournier, R.A. Kneeshaw, D., 2022. Effects of Viewing Geometry on Multispectral Lidar-Based Needle-Leaved Tree Species Identification. Remote Sensing 14 (24), 6217.

Liu, B., Huang, H., Su, Y., Chen, S., Li, Z., Chen, E. Tian, X., 2022. Tree species classification using ground-based LiDAR data by various point cloud deep learning methods. Remote Sensing 14 (22), 5733.

Lombard, L., Ismail, R. and Poona, N.K., 2020. Modelling forest species using Lidar-derived metrics of forest canopy gaps. South African Journal of Geomatics 9 (1), 31-43.

Novo, A., Gonzalez-Jorge, H., Comesana-Cebral, L.J., Lorenzo, H., Martinez-Sanchez, J., 2022. Semi-automated tree species classification based on roughness parameters using airborne lidar data. DYNA 97 (5), 528-534.

Verma, M.K. Yadav, M., 2024. Estimation of plant's morphological parameters using terrestrial laser scanning-based three-dimensional point cloud data. Remote Sensing Applications: Society and Environment 33, 101137.

Yadav, M., Khan, P., Singh, A.K. Lohani, B., 2021. An automatic hybrid method for ground filtering in mobile laser scanning data of various types of roadway environments. Automation in Construction 126, 103681.