Performance Assessment of Neural Radiance Fields (NeRF) and Photogrammetry for 3D Reconstruction of Man-Made and Natural Features

Abhinav Jagan Polimera, M. M. Prakash Mohan[a] and K. Rajitha[b]
Birla Intitute of Technology and Science Pilani, Hyderabad Campus, India

Keywords: Neural Radiance Fields (NeRF), Photogrammetry, 3D Reconstruction, Ecological Modeling.

Abstract: The present study focuses on the reconstruction of 3D models of an antenna (man-made) and a bush (natural feature) by adopting the recently developed Neural Radiance Fields (NeRF) technique of deep learning. The performance of the NeRF was compared with the outcomes obtained by the traditional photogrammetry methods. The ground truth geometric observation of the selected objects derived using electronic distance measurement-based techniques revealed the efficacy of NeRF compared to photogrammetry for both man-made and natural features' reconstruction cases. The capabilities of NeRF to reconstruct the features with complex geometries were evident from the outcome of bush 3D reconstruction. The prospectus of canopy and leaf level geometry estimation using NeRF will aid the enhanced modeling of vegetation-atmosphere interactions. The findings presented in the study have significant implications for diverse fields, from entertainment to ecological modeling, and offer insights into the practical applications of NeRF in 3D reconstruction. The outcomes of the present study attempted with a texture-less object like a bush unveiled the opportunities to apply the NeRF techniques in precision agriculture.

1 INTRODUCTION

Three-dimensional (3D) models are versatile tools with various applications across industries, from entertainment to education, healthcare to engineering. They enhance visualization, planning, and problem-solving by providing an immersive and interactive experience that traditional 2D representations cannot match. They offer a realistic and immersive way to visualize and interact with objects, spaces, and concepts. The scientific community, especially environmental researchers, considers the development of 3D reconstruction techniques as a boon to augment the ecological models with more structural attributes for model calibrations (Munier-Jolain et al., 2013).

The accurate modeling of heterogeneous features of the natural environment demanded a sophisticated data acquisition system to generate the 3D models. The high capital cost and processing requirements hampered the development of 3D models in the ecological domain and restricted them to 2D models, which created more gaps from real scenarios. The last decade's prominent focus on climate change-related research explored new ways of implementing 3D model-derived parameters adopting the advances in computer vision techniques. In the previous two decades, digital photogrammetry techniques revolutionized 3D topographic mapping with sufficient overlapping stereo-pairs(Chandler, 1999). The requirement of sufficient overlapping photos of high resolutions attenuated the traditional photogrammetry technique's applications in heterogeneous feature reconstruction.

The surge in artificial intelligence and machine learning techniques has catalyzed a revolution across numerous domains within science and engineering. The search for an appropriate technique that requires a minimum number of photographs for 3D reconstruction converged towards the recent approach of NeRF (Neural Radiance Field). The NeRF technology, supported by a complex neural network, enabled the rapid and accurate generation of 3D models (Palestini et al., 2023).

The present study targets to utilize the capabilities of NeRF for reconstructing the 3D image of a man-made feature and a natural heterogenous features. The research questions we address in this con-

[a] https://orcid.org/0000-0003-2484-4964
[b] https://orcid.org/0000-0003-2269-1933

Polimera, A., Mohan, M. and Rajitha, K.
Performance Assessment of Neural Radiance Fields (NeRF) and Photogrammetry for 3D Reconstruction of Man-Made and Natural Features.
DOI: 10.5220/0012396700003636

In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence (ICAART 2024) - Volume 3*, pages 840-847
ISBN: 978-989-758-680-4; ISSN: 2184-433X

text are how effectively the NeRF extracts the geometrical attributes of both man-made and natural features and how it differs from photogrammetry-derived outcomes. Using ground-verified geometrical parameters estimated by accurate electronic distance measurement techniques enabled the comparison between the performances of NeRF and photogrammetry outcomes.

2 RELATED WORK

Various global authors discussed the role of 3D data-driven spatial models in understanding different aspects of ecological modeling like forest changes, resilience to climate change, and capacity for carbon sequestration(Huang et al., 2019). Ecological modeling requires accurate canopy architecture quantification by determining the leaf area index (LAI). The reconstruction of 3D natural features like a plant canopy geometry remains challenging due to heterogeneous features and questions about the geometric fidelity of classical approaches (Xu et al., 2021). The research carried out in the recent decade related to 3D reconstruction of trees paid less attention to the estimations of geometrical parameters like canopy crown volume, tree height, vertical and horizontal distributions of foliage and leaf angle (Gromke et al., 2015). These geometric parameters are essential to understand the climate-related aspects of the ecosystem.

Two commonly applied techniques for 3D reconstruction of natural features involve active range data obtained through structured light sources such as lasers, and the other approach utilizes overlapping photos in conjunction with stereoscopic vision (Sequeira et al., 1999). The leaf level traits through these methods were less attempted due to high-cost in terms of data volume and processing requirements.The recent developments in deep-learning-based Neural Radiance Fields (NeRF) that focus on synthesizing new views of 3D objects and reconstructing 3D shapes from a collection of images pave the way towards better geometric estimations(Mildenhall et al., 2021).

NeRF represents a significant shift in 3D computer vision and has shown remarkable potential in generating novel views of complex scenes (Tancik et al., 2023). The prospect of NeRF signifies a change in research towards a more holistic understanding and modeling of three-dimensional scenes, especially for a natural environment. NeRF's ability to represent detailed scene geometry with complex occlusions makes it suitable for canopy architecture-related studies which is less attempted in the present stage of research. The fewer views requirement and effective capture of geometric features from heterogenous environments make the NeRF technology a better option for 3D reconstruction (Deng et al., 2022).

3 APPROACH

The overall methodology adopted for the present study is shown in Figure 1. In the present study, we have examined Neural Radiance Fields (NeRF) and photogrammetry, two methods used for 3D modeling and reconstruction. For the comparative analysis adopted in the study, two distinct objects were considered. The first object, an antenna, as seen in Figure 2 (b), is predominantly composed of precise geometrical shapes, while the second object chosen originates from nature, specifically, a bush (Figure 2 (a)). The rationale behind this selection lies in the aspiration to assess the effectiveness of both methodologies in diverse contexts. It can be asserted that the process of 3D reconstruction for geometrically flawless objects is inherently less complex when juxtaposed with the reconstruction of natural objects, which inherently feature a greater degree of irregularities on their surfaces.The assessment procedure hinges upon the utilization of RGB (Red, Green, Blue) images derived from a video source. These images, representing individual frames extracted from the video stream, serve as the foundational input for the evaluation process. The capabilities of NeRF and photogrammetry techniques for 3D reconstruction are discussed further, along with a comparison.

Photogrammetry is a versatile and widely used technique for creating 3D models or reconstructing objects and scenes from photographs. The overlapping images captured from different viewpoints serve as the input data for the reconstruction process. In the initial stages of photogrammetry, distinct features are identified and matched across overlapping images. These features could include points, lines, or other visually distinct elements. This matching process establishes the correspondence between the same feature in different images.

Accurate reconstruction in photogrammetry requires understanding the internal and external parameters of the cameras used to capture the images. The calibration of the camera's intrinsic properties, such as focal length and lens distortion, and determining its position in 3D space are required for accurate 3D reconstruction. Using the calibrated camera parameters and the correspondences established in the feature extraction step, 3D points, are reconstructed through triangulation. Triangulation is a mathematical process that estimates the 3D coordinates of the features by

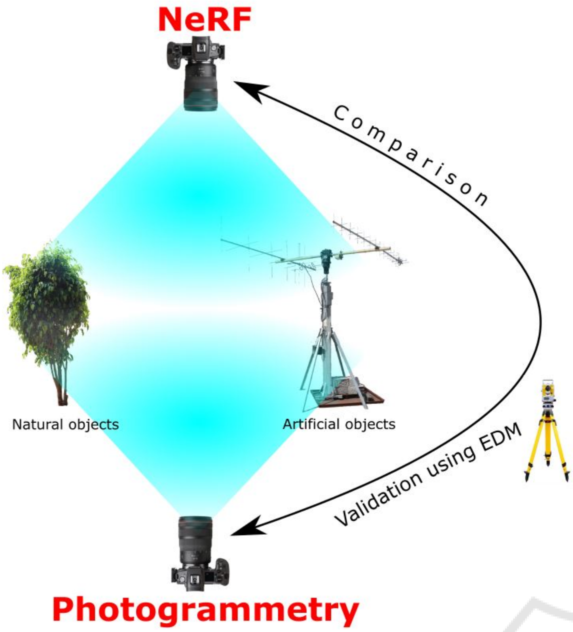

Figure 1: Flowchart of methodology.

determining where lines of sight from different camera positions intersect in 3D space. The reconstructed 3D points create a 3D surface or mesh.

The technical capabilities of NeRF make it a better option for the present study as a key differentiator compared to traditional neural networks for 3D reconstruction. Given a set of images capturing the same object from multiple angles along with their respective poses, NeRF learns to represent the 3D object in a way that enables the consistent synthesis of new views based on the training set. This instance-specific nature allows NeRF to model and represent the subtle, object-specific details.

The fundamental architecture of NeRF (Mildenhall et al., 2021) involves a simple Multilayer Perceptron (MLP). This MLP takes a single 5D coordinate as input, comprising three dimensions for location (x, y, z) and two for viewing direction (θ,Φ). The output of the MLP includes the density and color attributes at that spatial location. In practice, the location is mainly used to predict density, while viewing direction is combined with other information to predict color. Despite its simplicity, this basic architecture has demonstrated the ability to perform complex tasks and effectively capture scene geometry and appearance. A neural network is trained to estimate the color and intensity for each point in 3D space based on the input images and camera poses. This neural network learns a mapping from 3D coordinates and viewing directions to radiance values.

In NeRF, the synthesized views are obtained by

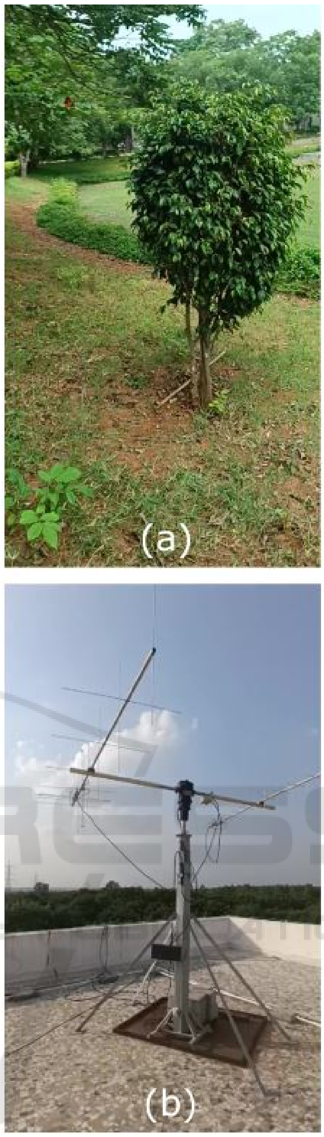

Figure 2: (a) bush (b) antenna.

querying 5D coordinates along camera rays. Classic volume rendering techniques are then employed to project the output colors and densities into a 2D image. The process is based on the input images and their known camera poses, and it allows for the generation of photorealistic novel views of scenes, even when dealing with complex geometry and diverse appearance.

Photogrammetry is a traditional method for 3D reconstruction, while NeRF is a deep learning-based approach. Photogrammetry involves capturing multiple 2D images of an object or scene from different angles, while NeRF takes a collection of 2D images and their corresponding camera poses as input. In photogrammetry, a process identifies common features in these images (e.g., points or edges). It uses triangulation techniques to determine the 3D position of these features. At the same time, in NeRF, a neural network is trained to model a continuous 3D scene representation by learning a function that maps 3D coordinates to RGB values.

In practical terms, NeRF is a more recent approach that leverages deep learning to create 3D models, excelling in complex and challenging scenes, especially those with unique lighting or reflective properties. However, it demands significant computational resources during training. In contrast, while being computationally intensive during reconstruction, photogrammetry is a well-established and versatile technique suitable for a wide range of scenarios.

The evaluation of both methodologies followed the extraction of images from a video acquired through circumferential movement around the objects. The frame extraction rate from the video was systematically varied, consequently altering the quantity of images employed for the 3D reconstruction process. Specifically, frame rates of 2, 3, 4, and 5 frames per second were assessed, and the resultant images were subsequently processed through the respective programs.

4 RESULTS

The results obtained after applying the photogrammetry and NeRF are discussed in this section, along with the corresponding reconstructed images of the antenna and the bush. We conducted a comparative analysis of the quantified metrics, encompassing the number of triangles, edges, and vertices, derived from the resulting meshes for various cases of input images. The outcomes from the 3D reconstruction for the antenna and bush are visually represented in Figure 3, Figure 4. The numerical values of the enumeration of triangles, vertices, and edges, are presented in Table 1 for bush and Table 2 for antenna for the 3D model obtained through NeRF.

Table 1: Number of images versus the geometric attributes for the natural object (bush)-NeRF.

Images	76	114	152	190
Triangles	3225700	3000000	2870000	2400000
Vertices	1630000	1500000	1490000	1200000
Edges	4850000	4475000	4373000	4320000

The outcomes of the comparative analysis, which relies on the enumeration of triangles, vertices, and edges, are conspicuously presented in Table 3 for bush and Table 4 for antenna for the 3D model obtained through photogrammetry, in addition to being visually represented through the accompanying graph.

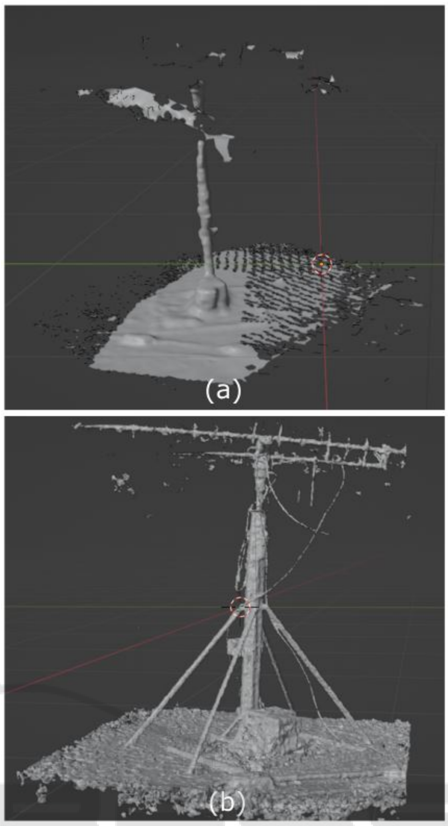

Figure 3: Man-made object (a) Photogrammetery result (b)NeRF result.

Table 2: Number of images versus the geometric attributes for the man-made object (antenna)-NeRF.

Images	76	114	152	190
Triangles	1260000	1240000	1186000	1060000
Vertices	636000	629000	6025600	581877
Edges	1890000	1870730	1785850	1146500

Table 3: Number of images versus the geometric attributes for the natural object (bush)-Photogrammetry.

Images	76	114	152	190
Triangles	52000	73000	86000	100000
Vertices	26000	36000	42000	58000
Edges	78000	110000	132000	140000

The outcomes of the comparative analysis, as delineated in Table 1 and the accompanying graph (Fig-

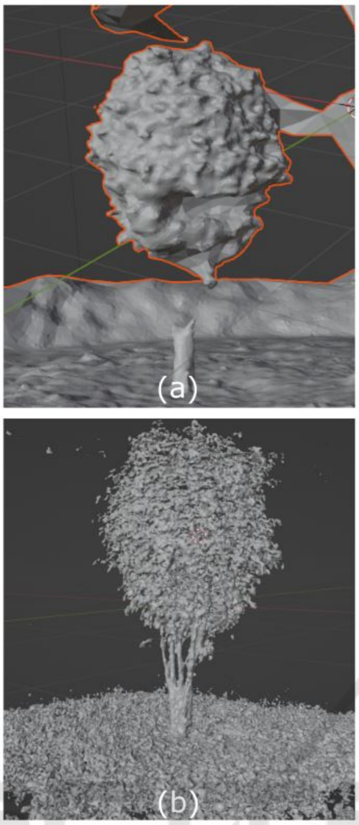

Figure 4: Natural object (a) Photogrammetery result (b) NeRF result.

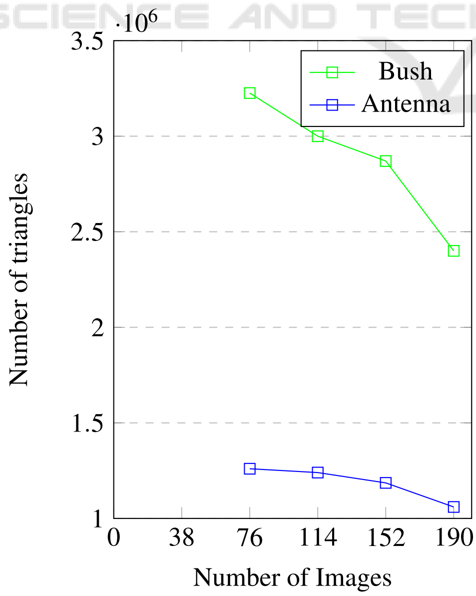

Figure 5: Plot of number of images versus number triangles-NerF.

Table 4: Number of images versus the geometric attributes for the natural object (antenna)-Photogrammetry.

Images	76	114	152	190
Triangles	32000	38000	50000	75000
Vertices	20000	23000	33000	42000
Edges	51000	58000	70000	95000

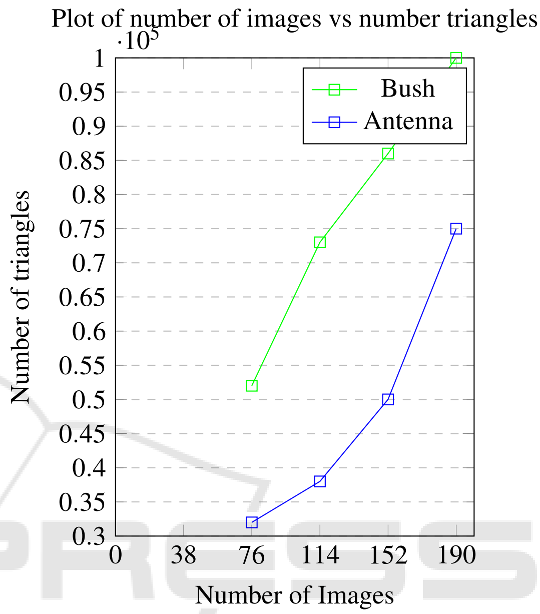

Figure 6: Plot of number of images versus number triangles-Photogrammetry.

ure 5), elucidate notable trends. The graph pertaining to NeRF (Neural Radiance Fields) reveals a discernible reduction in the number of triangles with an increase in the number of input images. This phenomenon can be attributed to NeRF's initial assumption of each data point as an independent vertex in three-dimensional space. However, as more images are incorporated, NeRF acquires a broader context and amalgamates points corresponding to the same surface. Consequently, this integration leads to a decrease in the count of vertices, triangles, and edges.

Conversely, the graph representing photogrammetry(Figure 6) exhibits an opposite trend, with an augmentation in its geometric complexity as more images are introduced. This phenomenon can be attributed to the extraction of additional feature points from each image, facilitating improved inter-image matching and increasing the degree of overlap among the acquired images. In light of the aforementioned findings, it is evident that the acquisition of geometric attributes, such as height, area, and volume, can be achieved even with a limited number of input images.

Table 5: Comparison of the geometrical attributes estimated by NeRF and Photogrammetry.

	Height (Man-made object-Antenna)			
	Modelled	Ground Truth	Error	Accuracy %
NeRF	2.993 m	2.968 m	0.025 m	99.157
Photogrammetry	3.021 m	2.968 m	0.053 m	98.21
	Height (Natural object-Bush)			
NeRF	1.794 m	1.792 m	0.002 m	99.88
Photogrammetry	1.81 m	1.792 m	0.018 m	98.99
	Max.Width (Natural object-Bush)			
NeRF	0.961 m	0.96 m	0.001 m	99.89
Photogrammetry	1.046 m	0.96 m	0.116 m	87.91

Specifically, our analysis focused on determining the height of the objects under consideration and subsequently comparing these measurements to ground truth values. To obtain precise object dimensions, we judiciously employed control points and proportionally scaled the reconstructed mesh. As a result, the height of the antenna, as derived from the NeRF model, was 2.993 m, while the height determined through the Photogrammetry model was 3.021 m, and the ground truth value was 2.968 m (Figure 7). This comparison revealed that NeRF yields results closer to the ground truth than photogrammetry. Similarly, when evaluating the dimensions of the bush, the respective length, breadth, and height parameters were very close to ground truth values in the case of NeRF compared to photogrammetry-derived results (Figure 8, Table 5). In the case of bush, the geometric values shown in Figure 8 are the average values of measurements taken in different directions across the canopy volume.

Performing an exact computational analysis presents significant technical challenges due to the fundamental differences in the underlying algorithms employed by NeRF (Neural Radiance Fields) and photogrammetry methods. NeRF offers the advantage of real-time view rendering. Conversely, photogrammetry directly produces the finalized rendered model without the intermediate step of rendering individual views.

Empirical observations underscore the efficiency discrepancy between NeRF and photogrammetry under favorable lighting conditions. NeRF demonstrates remarkable rapidity, accomplishing approximately most of the rendering process(sufficient for extracting critical geometric features) within a few seconds. In contrast, the output derived from photogrammetry takes considerably longer, spanning several minutes to complete the rendering process under similar conditions. This discrepancy in rendering times highlights the substantial disparity in computational efficiency between NeRF and photogrammetry methodologies, particularly in scenarios characterized by optimal lighting conditions.

The comprehensive analysis indicates a superior performance by NeRF across the evaluated criteria. Furthermore, the potential applications of NeRF extend to diverse domains, including the creation of canopy height models. An additional breakthrough lies in the precise determination of leaf angles using NeRF, as it consistently produces highly accurate leaf models, a feat not achieved as effectively by photogrammetric reconstructions.

5 CONCLUSIONS AND FUTURE RECOMMENDATIONS

The present study underscores the significance of 3D models and presents a comparative analysis of two 3D model reconstruction techniques, namely Neural Radiance Fields (NeRF) and photogrammetry. The key findings and implications derived from the study can be summarized as follows:

- Efficiency of NeRF: NeRF stands out as a highly efficient method for determining the geometric dimensions of objects due to its ability to achieve accurate results with a reduced number of input images. This efficiency is further emphasized by its computational speed, making it a practical choice for 3D reconstruction tasks.
- Accurate Complex Geometry Capture: The study demonstrates that NeRF excels in capturing complex geometries, as evidenced by the highly accurate 3D reconstruction of the bush model. This exceptional accuracy implies that NeRF holds significant potential for a wide range of applications

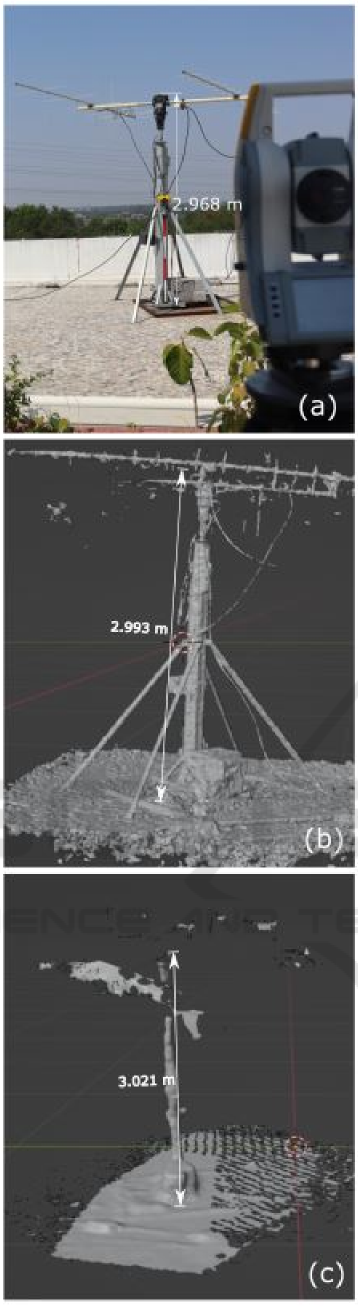

Figure 7: Validation of man-made object (a) Ground Truth (b) NeRF result (c) Photogrammetry result.

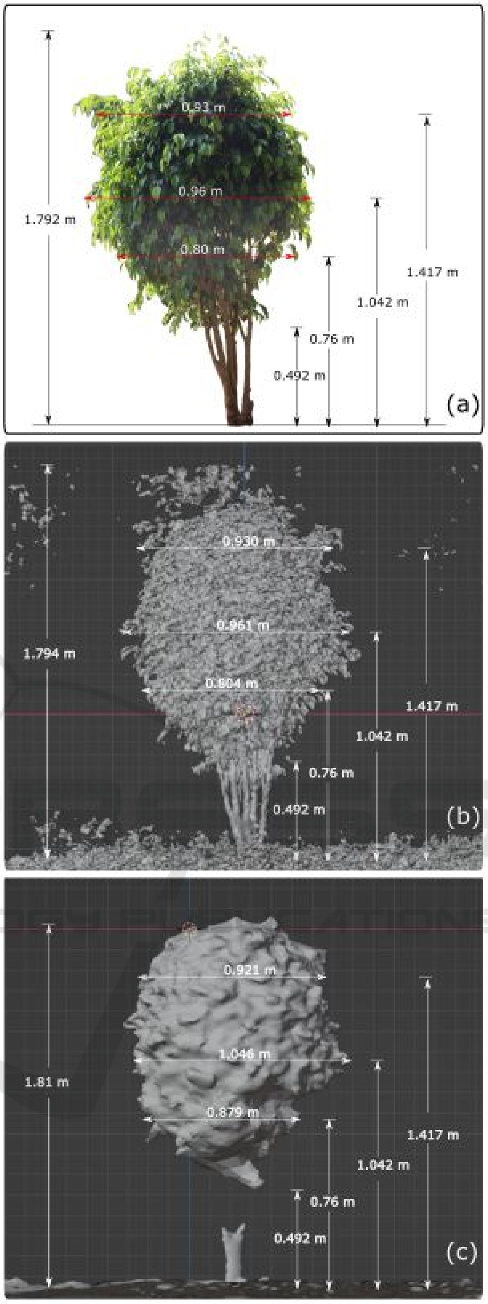

Figure 8: Validation of natural object (a) Ground Truth (b) NeRF result (c) Photogrammetry result.

where intricate geometries need to be faithfully represented.

- Leaf Angle Measurement: NeRF's proficiency in dealing with complex geometries extends to measuring leaf angles, a task that is challenging for photogrammetry due to the lack of sufficient detail in the models it generates. NeRF's ability to capture fine details makes it suitable for applications such as leaf angle measurement, which is critical in various contexts.
- Canopy Height Models: NeRF's efficient performance in determining the geometric dimensions of objects has practical implications in the creation of canopy height models for trees. This method allows for faster and more accurate canopy height modeling with a reduced number of

input images, streamlining the process of assessing tree canopies.

- Leaf Area Index estimations: The efficacy of NeRF revealed in the present study is directed towards its potential for estimating leaf area index, which is one of the essential climate variables. LAI is a crucial input for the various vegetation-atmosphere interaction models like production efficiency models for estimating the gross primary productivity of vegetation. Future research in this direction enables accurate estimation of energy fluxes, which is of prime importance in today's climate-changing world.
- The approach applied in the present study is well-aligned with the requirement of the high-throughput phenotyping (HTP) system for cash crops like cotton, which enables phenotypic trait estimation through these non-invasive 3D imaging techniques.

REFERENCES

Chandler, J. (1999). Effective application of automated digital photogrammetry for geomorphological research. *Earth surface processes and landforms*, 24(1):51–63.

Deng, K., Liu, A., Zhu, J.-Y., and Ramanan, D. (2022). Depth-supervised nerf: Fewer views and faster training for free. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12882–12891.

Gromke, C., Blocken, B., Janssen, W., Merema, B., van Hooff, T., and Timmermans, H. (2015). Cfd analysis of transpirational cooling by vegetation: Case study for specific meteorological conditions during a heat wave in arnhem, netherlands. *Building and environment*, 83:11–26.

Huang, J., Lucash, M. S., Scheller, R. M., and Klippel, A. (2019). Visualizing ecological data in virtual reality. In *2019 IEEE conference on virtual reality and 3D user interfaces (VR)*, pages 1311–1312. IEEE.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R. (2021). Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1):99–106.

Munier-Jolain, N., Guyot, S., and Colbach, N. (2013). A 3d model for light interception in heterogeneous crop: weed canopies: model structure and evaluation. *Ecological modelling*, 250:101–110.

Palestini, C., Meschini, A., Perticarini, M., and Basso, A. (2023). Neural networks as an alternative to photogrammetry. using instant nerf and volumetric rendering. In *Beyond Digital Representation: Advanced Experiences in AR and AI for Cultural Heritage and Innovative Design*, pages 471–482. Springer.

Sequeira, V., Ng, K., Wolfart, E., Gonçalves, J. G., and Hogg, D. (1999). Automated reconstruction of 3d models from real environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(1):1–22.

Tancik, M., Weber, E., Ng, E., Li, R., Yi, B., Wang, T., Kristoffersen, A., Austin, J., Salahi, K., Ahuja, A., et al. (2023). Nerfstudio: A modular framework for neural radiance field development. In *ACM SIGGRAPH 2023 Conference Proceedings*, pages 1–12.

Xu, H., Wang, C. C., Shen, X., and Zlatanova, S. (2021). 3d tree reconstruction in support of urban microclimate simulation: a comprehensive literature review. *Buildings*, 11(9):417.