



ORIGINAL ARTICLE OPEN ACCESS

3D Reconstruction of Small Flexible Objects With Slender Structures: Reconstructing Conifer Seedlings for Development of Computer Vision Systems in Virtual Environments

Elias Pääkkölä | Håkan Lideskog | Magnus Karlberg

Machine Design, Luleå University of Technology, Luleå, Sweden

Correspondence: Elias Pääkkölä (elias.paakkola@ltu.se)

Received: 17 September 2025 | **Revised:** 25 May 2026 | **Accepted:** 27 May 2026

Keywords: 3D reconstruction | 3D scanning | automated workflow | close-range photogrammetry | digitization | image-based modeling

ABSTRACT

3D reconstruction has matured into a robust technology. However, small, flexible objects such as conifer seedlings remain challenging due to their fine-scale structures, and susceptibility to movement. This study investigates and evaluates methods for reconstructing spruce (*Picea abies*) and pine (*Pinus sylvestris*) seedlings, with the aim of establishing a workflow capable of capturing geometry and texture for applications in machine learning and virtual testing environments. Two acquisition approaches were tested: photogrammetry using a RGB camera and a 3D scanner, both mounted on a robotic arm. While the scanner produced incomplete results, the photogrammetry approach successfully generated point clouds (pcl) with color information. Three different photogrammetry software were tested before relying on Agisoft Metashape and Meshroom for image processing and dense pcl generation, followed by pcl filtering in CloudCompare and meshing in Blender. Six seedlings were reconstructed to textured meshes and quantitatively evaluated using the metrics precision, recall, F1-score, mask intersection-over-union (IoU), and boundary IoU. Results showed an average mask IoU of 75.7% and F1-score of 86.1%. Pine seedlings yielded higher recall and F1-scores, whereas spruce reconstructions demonstrated higher precision. The proposed semi-automated workflow demonstrates the feasibility of reconstructing small and slender structured flexible objects, specifically conifer seedlings.

1 | Introduction

3D reconstruction has rapidly evolved into a mature and widely applied technology, driven by advances in sensors, algorithms, and computational power (Ham and Wesley 2019). Today, it underpins applications ranging from cultural heritage documentation to industrial inspection, medical imaging, and environmental monitoring. Established solutions and automated workflows for 3D reconstruction (Remondino et al. 2017; Galantucci et al. 2018; De Paolis et al. 2020; Yang et al. 2024), demonstrate the accessibility and robustness of modern reconstruction pipelines, while recent research explores automation and reconstruction through artificial intelligence (Abbood

et al. 2024; Hu et al. 2024), including emerging deep learning-based approaches such as Neural Radiance Fields (NeRF), which enable photorealistic reconstruction from multi-view imagery (Hu et al. 2024), but are still in its inception regarding mesh construction for virtual environments (Frasiolas et al. 2025).

Beyond established domains, virtual environments and digital twins increasingly rely on accurate 3D representations of real-world objects and environments. Such reconstructions are critical for training neural networks, simulating machine-object interactions, and testing autonomous systems in controlled, repeatable, and safe conditions (Lehto et al. 2026). Pioneering this transformation is the automotive industry, where prototyping

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *The Photogrammetric Record* published by Remote Sensing and Photogrammetry Society and John Wiley & Sons Ltd.

and simulation have reduced development time and physical testing (Schmidt et al. 2023). In recent years, these simulators have been used to test autonomous systems in diverse and repeatable conditions, while enjoying the advantage of testing scenarios where it otherwise may be dull, dirty, or dangerous for operators (Dosovitskiy et al. 2017). In forestry, this development is particularly relevant. Some operations are still mainly performed manually, especially in forest regeneration, despite ongoing efforts toward automation. The forest environment is inherently sensitive, heterogeneous, and complex, where small, flexible objects play a key role in the interaction between human, machine, and the ecosystem. Although highly mechanized and automated systems exist, exemplified in the Autoplant project, the BraSatt initiative, and the PlantMaX machine, they still require robust solutions that can replace human capabilities while ensuring high-quality outcomes (Hansson et al. 2024). Accurate 3D representations of objects could therefore enable realistic simulations of forest machine-environment interactions and transfer costly and destructive testing to a virtual setting (Lehto et al. 2026).

Reconstructing small forest objects is challenging though. Their fine-scale, repetitive structures and sensitivity to movement complicate both acquisition and processing, as noted in plant phenotyping and forest monitoring studies (Yin et al. 2024; Cova et al. 2023). Nevertheless, successful reconstruction would enable new opportunities: the generation of synthetic variants for auto-annotated training datasets, the integration of realistic seedling models into full-scale machine simulations, and the testing of control algorithms without causing physical damage to soil, plants or equipment. 3D reconstruction of tree seedlings has been done, but on seedlings with the height between 20 and 200 cm and excluding (not mentioning) conifer trees (Yang et al. 2022). Pisek et al. (2023) have successfully reconstructed the geometry of conifer shoots but with a methodology that did not retain color and texture, which is an essential feature when training image based computer vision model (Krizhevsky et al. 2017), prevalent in most automation efforts (LeCun et al. 2015). Meanwhile, James et al. (2023) reconstructed plants successfully with color and texture but encountered difficulties with slender features. While a lot of research has been done on reconstructing small objects with thin features in one dimension, e.g., different broad-leaved plants, there is still a research gap regarding how to reconstruct small flexible objects with thin features in two dimensions, such as conifer seedlings, while also enabling their use in synthetic environments for computer vision model training.

Hence, the purpose of this paper is to enable utilization of realistic virtual environments, for applications in, e.g., forestry automation (computer vision), machine learning, and virtual testing environments. The goal is to establish a workflow for creating digital representations of the geometry and texture of small and slender structured flexible objects with special emphasis on conifer seedlings that are used in Nordic silviculture operations.

The main contributions of this work can be summarized as follows:

- A semi-automated reproducible workflow for 3D reconstruction of small, slender, and flexible objects, dedicated to computer vision purposes.

- an evaluation methodology suitable for quality assessment of slender object reconstruction, and
- a dataset with raw image data and finalized textured meshes of six conifer seedlings, commonly used in Nordic forest regeneration.

1.1 | Related Research

Over the past decades, significant research and development in 3D scanning and reconstruction have been driven by technological advancements in both hardware and software, supported by the rapid increase in computational power. Ham and Wesley (2019) presents a brief overview of the development of 3D reconstruction spanning from 1981 to 2018 and claims the field to be mature considering well established commercial products. In recent years, effort has been made to automate 3D-reconstruction. Remondino et al. (2017) presents an overview of automated photogrammetric processing of large datasets and the performance of commercial software packages, focusing on reconstruction of cultural heritage objects and scenes, while Yang et al. (2024) and Stausberg et al. (2024) automatically 3D reconstructs broad-leaved plants, although the later does not include the image acquisition in the automated pipeline. Galantucci et al. (2018) published a review of the current state of photogrammetry applied to small and micro-scaled objects with sub-millimeter precision, including scanners, algorithms and methodologies with some attention to robustness and automation. They claim that this strategy is a good option to contact measurements under certain conditions. De Paolis et al. (2020) conducted a survey on photogrammetry on small objects covering all phases of image acquisition, 3D reconstruction and texture extraction, including automation, hardware setup, algorithms and software. When it comes to 3D modeling and reconstruction of plants and trees, it has been reviewed by Okura (2022), highlighting difficulties with occlusion and hidden structures. Reconstructing forest scenes and objects, structure-from-motion (SFM) is one common technique reviewed by Iglhaut et al. (2019), while Magnuson et al. (2024) conducted a more recent review covering not only SFM that examines mobile devices, such as phones and tablets, as measurement devices. Both reviews, however, only cover bigger objects such as trees. Cova et al. (2023), evaluated SFM close-range photogrammetry for understory forest biomass and found promising results on all levels except biomass in the height range of 0-10 cm. As mentioned, it is a challenging task to reconstruct unstructured objects that have repetitive and homogenic features, and the complexity increases with objects that are small and not rigid or fixed, such as plants and insects/animals, which for example is shown by studies (James et al. 2023; Shamblin and Puppo 2024). Yin et al. (2024) has made a comprehensive review of 3D reconstruction and image acquisition methods based on monocular stereovision applied on plants. Pisek et al. (2023) conducted blue-light photogrammetry scanning on pine- and spruce shoots (not seedlings), successfully reconstructing their geometry and demonstrating a promising workflow for capturing small and slender flexible objects. Robotics have been used to successfully automate the process of image acquisition, such as 3D reconstruction of indoor environments (Mortezapoor et al. 2022), robot arm utilization applied on plant phenotyping (Karty and Hament 2024) and insect reconstruction (Muthu

et al. 2023), while a lot has been done with automatic rotary tables, e.g., (Plum and Labonte 2021; Sirotin and Polyakov 2022). Using a drone for collection of UAV imagery (Marks et al. 2022) 3D reconstructed broad-leaved crops and Medina et al. (2020) 3D reconstructed living birds with submillimeter precision, using a setup of six cameras mounted on a rotary base. Noted by Yang et al. (2024) is that rotary tables often introduce jitter affecting the output quality of the image acquisition step.

2 | Methods and Materials

As noted in related research, small flexible objects are inherently difficult to reconstruct in 3D due to the scale and their susceptibility to movement and deformation. To address these challenges, our workflow was designed to ensure that the object remained completely at rest throughout the data capture process. The fundamental requirements for achieving this included a stable laboratory environment, avoidance of any physical movement of the object, and reduction of airflow or vibrations.

To develop a suitable methodology, we adopted a bottom-up strategy. A representative object was first selected for initial testing, and various approaches were evaluated through iterative trial and error. Methods were progressively refined and combined until a working workflow was established.

The first step involved identifying an appropriate data acquisition device and a capture strategy that preserved object stability. Once image or scan data were obtained, the next stage focused on reconstructing the object geometry. Depending on the chosen approach, additional intermediate steps were required before reaching the final reconstruction, such as cleaning and filtering raw point cloud data and converting them into textured meshes. While all examined approaches and processing stages are not explained in exhaustive detail here, they were distilled into a set of core approaches that are explored further in this article. Two data acquisition approaches were tested: (1) photogrammetry using a digital single-lens reflex (DSLR) camera, and (2) a 3D scanning approach using the same experimental setup mounted on a robotic arm. Three photogrammetry software packages

were then compared in terms of their reconstruction performance. Initial experiments were conducted on a single spruce (*Picea abies*) seedling, whereas the final workflow was validated on two additional spruce seedlings and three pine (*Pinus sylvestris*) seedlings, all provided by Svenska Skogsplanter AB (see Appendix B: Figure B.1 for images). The height of all seedlings was in the range of 6 to 15 cm (excluding the root ball).

The following section provides a detailed description of the developed workflow and its main components.

2.1 | Experimental Setup for Data Acquisition

The data acquisition was conducted in an environment-controlled room illuminated by four roof mounted fluorescent fixtures. Figure 1 shows the setup of the robot arm UR3e (Universal robots A/S, Odense, Denmark) which is surrounded by white sheets from all sides except from above where translucent white plastic boards are placed to soften up the lighting. This provisory setting has the function of a light box providing diffused lighting conditions to avoid shadows. The robot arm is placed in the middle of the light box standing on a footing to provide room for movement beneath the robot base. A teach pendant is used to control the robot arm and both are connected with separate cables to a control box. Attached one at a time, switched depending on the case, to the robot's end effector are two different 3D printed tool mounts. One mount (including a short wooden plank) for the 24 MP Nikon D5200 (Nikon Corporation, Nishioi, Shinagawa-ku, Tokyo, Japan) DSLR camera with the camera lens Nikkor 24-120 mm and one mount for the EinScan HX handheld 3D scanner (Shining 3D, Hangzhou, China), see Figure 1a,b for the respective setups. The seedling is attached on a rack hung from above to ensure full range of movement for the robotic arm. Two threaded rods attached to the top of the light box hold the rack. As shown in Figure 1c, the rack consists of a white bottom plate of cardboard, four wooden sticks and a blue 3D printed top frame. Markers are placed in an irregular pattern covering the entire rack. In the center of the bottom plate a short wooden stick holds the seedling.

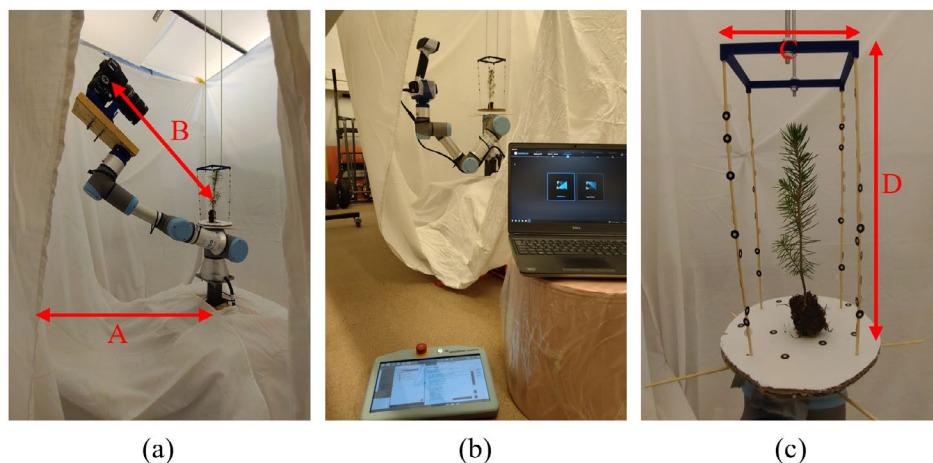


FIGURE 1 | Data acquiring setup. The experimental setup, to the left (a), is the DSLR camera setup; in the middle (b) is the 3D scanner setup, and to the right (c) is the cage holding the seedling. The distance between the sheets and the robot base is $A = 70$ cm, the distance between the camera base and the center of the cage is $B = 63$ cm, the width and height of the cage is $C = 10$ cm and $D = 31$ cm.

2.2 | Data Acquiring Procedure

The procedure for acquiring data differs between the scanner types, although we are minimizing the difference. The same robot arm was used in both cases and in this study the robot is controlled via the teach pendant manually or programmed with the inbuilt python-based URScript programming language. In Figure 2 the coordinate systems of the robot arm are illustrated. The end effector's position (origo of the denoted coordinate system) is defined by the modified spherical coordinates:

$$x = R \cos(\theta) \sin(90 - \varphi) \quad (1)$$

$$y = R \sin(\theta) \sin(90 - \varphi) \quad (2)$$

$$z = R \cos(90 - \varphi) \quad (3)$$

where θ is the rotation around the vertical (Z) axis and φ is the rotation around the horizontal plane. In these experiments θ ranges from 0° to 360° (one revolution) in steps of 5° and φ ranges from 0° to 30° in steps of 10° , which are in total four revolutions on different heights. Between each revolution of θ , φ was changed and to not reach the robot joints limits (defined as $\pm 360^\circ$) the rotation direction was changed as well. The end effector orientation is continuously changed to keep the scanner pointing at the same spot at the center of the experimental setup (center of seedling), i.e., changing μ matching the change of φ . In accordance to previous research such as (Yang et al. 2024; Róg and Rzonca 2021), the angular step sizes were selected to ensure

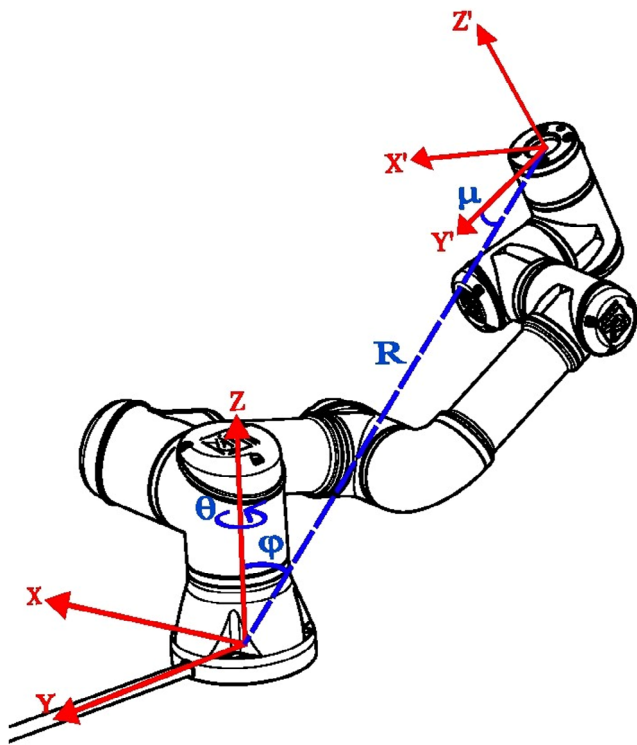


FIGURE 2 | Robot coordinate system. A schematic diagram of the base coordinate system (x, y, z) of the robot arm and the end effector coordinate system (x', y', z') related to each other with spherical coordinates (θ, φ, R). When $\theta = 0^\circ$ and $\varphi = 90^\circ$, $R = 0$ the denoted coordinate system is aligned with the base coordinate system. The $y'z'$ -plane is always kept coincident with the Z-axis.

an 80%–90% image overlap. The maximum φ -value was further limited to the physical setup.

2.2.1 | DSLR Camera Approach

The DSLR camera used in the experiments is equipped with a polarizing filter (Dörr GmbH, Neu-Ulm, Germany) and is attached vertically to the robot arm end effector with the optical axis parallel to the y' -axis and directed at a point situated on the Z-axis of the base coordinate system, corresponding to the center of the seedling being photographed, as shown in Figure 1a.

The image acquisition of each seedling is automated by using URScript to program the robot arm and the inbuilt camera option interval shooting. The camera is started first and thereafter the robot program ensuring that the interval shooting is synchronized to when the robot is still. The camera takes a photo with a 2 s shutter speed every fourth second synchronizing with the robot that moves during 1 s and stands still for 3 s. Photos are taken in steps of five degrees in the horizontal plane at four different heights adding up to in total 288 photos. The camera parameters used are presented in Table 1.

2.2.2 | 3D Scanner Approach

The 3D scanning device used in the experiments is an EinScan HX handheld 3D scanner (Shining 3D, Hangzhou, China) with a hybrid blue laser & led light source. The scanner is attached upright to the robot arm end effector with the field of view center directed at a point along the Z-axis of the end effector corresponding to the center of the scanned seedling, as seen in Figure 1b. The externally powered scanner is connected to a computer running the program used to control and monitor the ongoing scanning progress in real time. After the scanning is completed a point cloud is generated which can be saved and exported, but it is also possible to do post processing directly in the software and optional mesh generation. In this work the software scan workflow, defined by the manufacturer and presented in Figure 3, is followed to generate the point cloud and mesh of the scanned seedling.

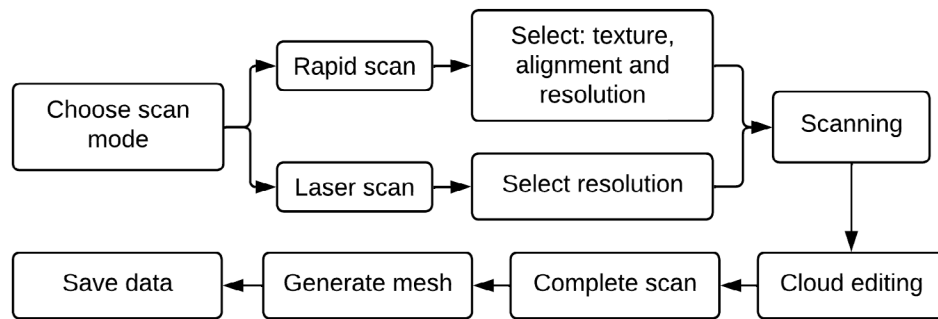
The scanning of each seedling is semi-automated by using URScript to program the robot arm. The predefined route described in Section 2.2 is performed in a continuous movement during several minutes (< 5 min). When finished, the scanned data is inspected in the real time software GUI, and areas with missing data are scanned by manually operating the robot arm.

2.3 | Photogrammetry Based 3D Reconstruction

After acquiring the RGB image data of the seedlings, it is uploaded to photogrammetry software. In the first stage, reconstruction of one spruce is performed in three different software: Agisoft Metashape 1.5.5.9097, RealityCapture 1.4.2 (both commercial), as well as the open-source software Meshroom v2.13.2. Agisoft was run on a different computer compared to the other two software; see Appendix A: Table A.1–A.3, Figure A.1 for hardware specifications. Following the analysis of the reconstruction results, since RealityCapture failed

TABLE 1 | The camera and 3D scanner parameters used during data acquisition.

Camera parameters		
Focal length (zoom)		35 mm
ISO		125
Aperture (<i>f</i> -value)		<i>f</i> /25
Shutter speed		2 s
Focus		Manual
Exposure		−0.7 (Auto)
Image quality		Jpeg 6000×4000, pixel size 6×6 μm
3D scanner parameters	Laser scan	Rapid scan
Light source	Laser	LED
Resolution	0.05–3 mm	0.25~3 mm
Accuracy	Up to 0.04 mm	Up to 0.05 mm
Alignment	Markers	Markers/features/hybrid/texture
Scanning speed	480,000 points/s, 55 FPS	1,200,000 points/s, 20 FPS
Texture Scan	No	Yes

**FIGURE 3** | EinScan HX workflow. The software scan workflow of the EinScan HX. Based on Shining 3D (2021).

to produce satisfactory 3D reconstruction (see Section 3.1 for details), Agisoft and Meshroom were selected as the tools for further processing.

All these photogrammetry packages share general theoretical frameworks (covered in depth by many sources, e.g., Luhmann et al. 2014; Ma et al. 2012; Hartley and Zisserman 2003), and are based on the same approach, i.e., extraction of image features, matching extracted features between images, creating structure-from-motion (SfM), meshing and mesh filtering.

Each image is modeled as a central projection (pinhole) camera with lens distortion. A 3D point $\mathbf{x} = (x, y, z)^T$ in object space is first transformed into the camera coordinate system:

$$\mathbf{x}_c = \mathbf{R}\mathbf{x} + \mathbf{t}, \quad (4)$$

where $\mathbf{R} \in \text{SO}(3)$ is the rotation matrix and $\mathbf{t} \in \mathbb{R}^3$ is the translation vector defining the exterior orientation.

Normalized image coordinates are obtained by perspective division:

$$\bar{x} = \frac{x_c}{z_c}, \quad \bar{y} = \frac{y_c}{z_c}. \quad (5)$$

Lens distortion is then applied using a radial-tangential model:

$$x_d = \bar{x}(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + 2p_1 \bar{x}\bar{y} + p_2(r^2 + 2\bar{x}^2), \quad (6)$$

$$y_d = \bar{y}(1 + k_1 r^2 + k_2 r^4 + k_3 r^6) + p_1(r^2 + 2\bar{y}^2) + 2p_2 \bar{x}\bar{y}, \quad (7)$$

where $r^2 = \bar{x}^2 + \bar{y}^2$, k_1, k_2, k_3 are radial distortion coefficients, and p_1, p_2 are tangential distortion coefficients.

Finally, the distorted coordinates are linearly mapped to pixel coordinates using the matrix:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & s & c_x \\ 0 & f(1+m) & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix}, \quad (8)$$

where f is the principal distance (focal length), (c_x, c_y) is the principal point, s is the skew parameter, and m represents differential scaling between the image axes.

Feature correspondences between images are established by detecting key points and matching local descriptors across image pairs. Matches are accepted based on descriptor similarity and

geometric consistency, typically enforced using the epipolar constraint.

The camera parameters and 3D point coordinates are estimated jointly using bundle adjustment, which minimizes the reprojection error over all image observations:

$$\min_{\{\mathbf{R}_i, \mathbf{t}_i, \theta_i, \{\mathbf{x}_j\}\}} \sum_{i,j} \rho(\|\mathbf{u}_{ij} - \hat{\mathbf{u}}_{ij}(\mathbf{R}_i, \mathbf{t}_i, \theta_i, \mathbf{x}_j)\|^2), \quad (9)$$

where $\mathbf{u}_{ij} = (u_{ij}, v_{ij})^\top$ are the observed image coordinates of point j in image i , and $\hat{\mathbf{u}}_{ij}$ are the corresponding projected coordinates obtained from the camera model described above. The parameter vector $\theta_i = (f, c_x, c_y, s, m, k_1, k_2, k_3, p_1, p_2)$ contains the intrinsic and distortion parameters of camera i , and ρ is a robust loss function used to reduce the influence of outliers.

The theory assumes a rigid scene, static camera intrinsics for each image, and perspective projection, with observations affected only by measurement noise and lens distortion. Hence, this theory corresponds to the established collinearity equations extended with lens distortion and constitutes the standard optimization problem underlying Structure-from-Motion (SfM) and Multi-View Stereo (MVS) pipelines.

Dense reconstruction by MVS then uses the estimated cameras to infer per-view depth maps and fuse them into a surface or dense point cloud. In MVS, an image/geometry explanation that is photometrically and geometrically consistent across views is typically searched for. Practical pipelines often compute depth maps first and then merge them into a dense surface representation.

The workflow in Agisoft started with uploading the images for each seedling in separate chunks. Next, camera alignment was conducted, creating a sparse point cloud and a reconstruction region for each chunk. This reconstruction region defines the volume of interest in the next step of calculating a dense point cloud and was manually fitted only to the plant and cage for each chunk, removing unnecessary background points and thereby reducing the needed computational power and time for computing the dense point cloud. The final step was to start a batch process that involves generating the dense point cloud and exporting the results for each chunk. The selected running parameters can be found in Appendix A but are mainly kept at default values other than quality (changed to high) and filter (turned off) settings. For an in-depth guide of the general SfM workflow in Agisoft (version 1.6) see (Over et al. 2021).

The workflow in Meshroom started with uploading the images of one seedling and then starting the photogrammetry pipeline. This procedure was then repeated for each seedling. The pipeline is built of several nodes; initially, camera intrinsic parameters are estimated, followed by feature extraction and matching to establish correspondences between images. Camera poses and a sparse point cloud were then reconstructed using SfM. Subsequently, the scene was prepared for dense reconstruction, and depth maps were computed for each image using multi-view stereo. These depth maps were filtered and fused to generate a dense point cloud that could be exported. See (Griwodz et al. 2021) for a detailed description of the default

photogrammetry pipeline. The running parameters that were changed from default can further be found in Appendix A.

After acquiring the dense point clouds, they were imported to the open source point cloud processing software CloudCompare v2, where two steps of processing were applied. The first step was manually removing the cage and background points from the plant using the segmenting tool. The second step was filtering RGB color, where the limit span was chosen empirically to [0–80 0–80 0–80], and points outside of that span were removed; see Section 2.3.1 for a sensitivity analysis.

The filtered point clouds were thereafter imported to the open source 3D creation software Blender 4.2 where a node-based geometric pipeline comprising two custom modifiers was applied to create the final mesh. The first modifier turns points into mesh and the second modifier removes loose vertices, edges and small mesh islands. The modifiers comprised of the following steps:

Point cloud conditioning: Outliers are removed based on nearest-neighbor distances, where points are discarded if the distance is larger than a user-defined threshold. Redundant points are merged within a user-defined radius enforcing a minimum spatial sampling density.

Surface reconstruction: The point cloud is converted to a mesh using an implicit volumetric approach, where a scalar field is constructed from the point set and the surface is extracted as an iso-surface. The voxel size controls resolution, while the iso-value governs surface smoothness.

Attribute transfer: Color information is assigned to mesh vertices using nearest-neighbor interpolation from the point cloud.

Geometric regularization: Mesh smoothing is applied using a Laplacian operator, reducing high-frequency noise while preserving overall shape.

Topological cleaning: Small disconnected components, long edges, and poorly connected vertices are removed based on user-defined threshold criteria, ensuring a consistent and well-defined mesh structure.

Projection refinement: Mesh vertices are projected onto the nearest points in the original point cloud to improve geometric fidelity.

The values used in the modifiers differ between spruce and pine. The detailed pipeline is found in Appendix C and descriptions of each node can be found in the Blender manual (Blender Documentation Team, n.d.) and source code.

2.3.1 | Reconstruction Quality Assessment

To be able to measure the quality of the final meshes a method is defined. Of each mesh four binary images are rendered in Blender which are then compared to ground truth data. The four camera positions are at $\varphi = 10^\circ$ and at four θ positions approximately 90° apart, covering all sides of the mesh. Each camera

position has a corresponding ground truth image, i.e., the original photograph of the seedling. The ground truth RGB image is converted to a binary mask with which the mesh image is compared. To separate the seedling from the background an edge detection is performed, primarily through a canny filter, but with manual inspection and editing. The threshold for creating the binary mask is chosen empirically and is subject to annotator bias. This threshold could further induce an error of up to a half percentage point of total area, but it is kept constant for all evaluations. The approximated error also includes the annotator bias in filling gaps where the edge detection of the seedling fails. A common machine learning set of metrics suited for pixel-to-pixel comparison is used, i.e., precision, recall, F_1 -score and mask intersection-over-union (IoU) (Rice et al. 2025). These metrics are calculated on the basis of True Positives (TP), False Positives (FP) and False Negatives (FN). In addition, Boundary IoU, introduced by (Cheng et al. 2021) to improve shape/boundary quality evaluation, is used. Figure 4 shows a color representation of the incoming values to the calculations of the mask IoU, precision, recall and F_1 -score (image a) and the boundary IoU (image b). In the zoomed section of image b the importance of choosing the right dilation for the calculation of the boundary IoU is highlighted. In this work 9 px were chosen empirically to fully cover the needle width, i.e., on average half of the measured width of several pine needles (pine is used since their needles are broader than those of spruce), and only the stem boundary (cf. Figure 4a,b).

2.3.2 | Sensitivity Analysis Color Filter

To derive a suitable color filter setting, a sensitivity analysis is performed (See Figure 5). The filter upper threshold for the RGB values is varied in steps of five between 80 ± 10 , keeping R, G and B the same at all times. Figure 5 shows that the average values for mask IoU and F_1 -score reach a plateau at 80 (still increasing slightly for 85 and 90) while the boundary IoU has a maximum at 80. The recall increases for each increase of filter threshold while the precision decreases demonstrating the relationship between these two measures and the trade-off one has to make when deciding this threshold, despite the F_1 -score staying stable.

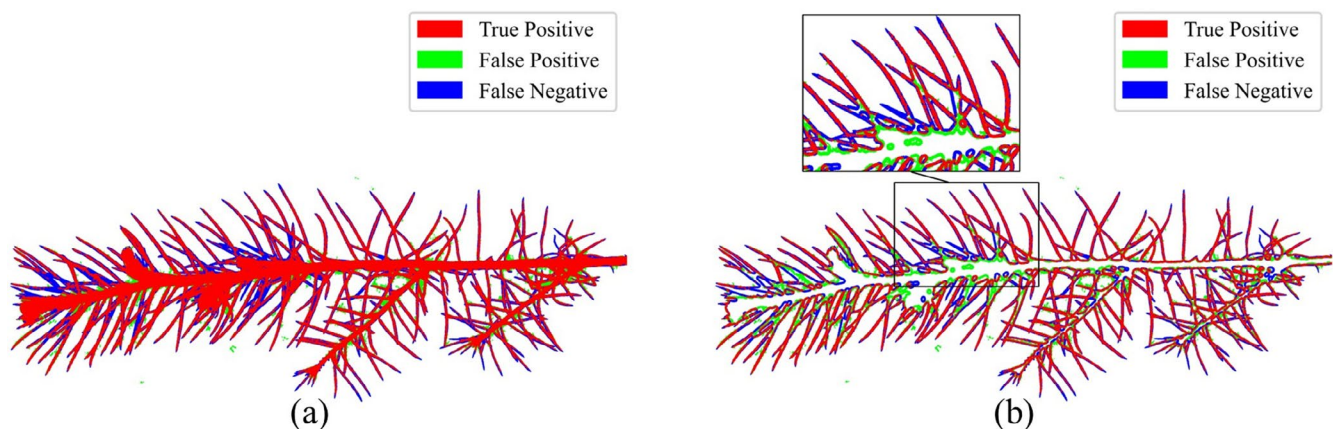


FIGURE 4 | Mask and boundary IoU. To the left (a) a color representation of the calculation basis of the mask IoU and to the right (b) how the boundary IoU is calculated. The red color represents the intersection of the ground truth and the prediction (reconstruction) called True Positive, the blue color represents only ground truth called False Negative and the green color represents only the prediction called False Positive. The zoomed area visualizes the chosen dilation from the boundary.

The number of points after applying the color filter decreases linearly from 1.4 to 1.2 million for thresholds between 90 and 70.

3 | Results

3.1 | Seedling 3D Reconstruction

The resulting dense point clouds for the test seedling (Spruce 1) generated in the three different photogrammetry software, Agisoft, Meshroom and RealityCapture, are presented in Figure 6, together with the resulting point cloud from the 3D scanning. In Table 2 the statistics of each software are presented, showing that Agisoft generated most points followed by the 3D scanning, RealityCapture and Meshroom. Agisoft was the most time-consuming and computationally heavy approach. However, it should be noted that the times are not directly comparable since Agisoft was run on another computer than Meshroom and RealityCapture; the hardware specifications can be found in Appendix A. All photogrammetry software had 100% camera alignment. After qualitative assessment of the point clouds, the decision was to move on with using photogrammetry for the remaining seedlings. As both the 3D scanning and RealityCapture clearly failed to capture much of the needles, they were excluded from further work. It is noteworthy to mention that RealityCapture only provides the dense point cloud corresponding to the generated mesh, resulting in data loss due to extensive filtering.

Table 3 summarizes the number of points and mesh elements throughout the processing pipeline. The initial photogrammetric point clouds differ substantially between Agisoft and Meshroom, where Agisoft generally produces denser point clouds (approximately 14–23 million points) compared to Meshroom (approximately 1.6–4.7 million points).

Following point cloud filtering, a substantial reduction in the number of points is observed for both methods. The first filtering step (manual cropping) removes a large portion of background points, while subsequent color filtering further reduces the point cloud size.

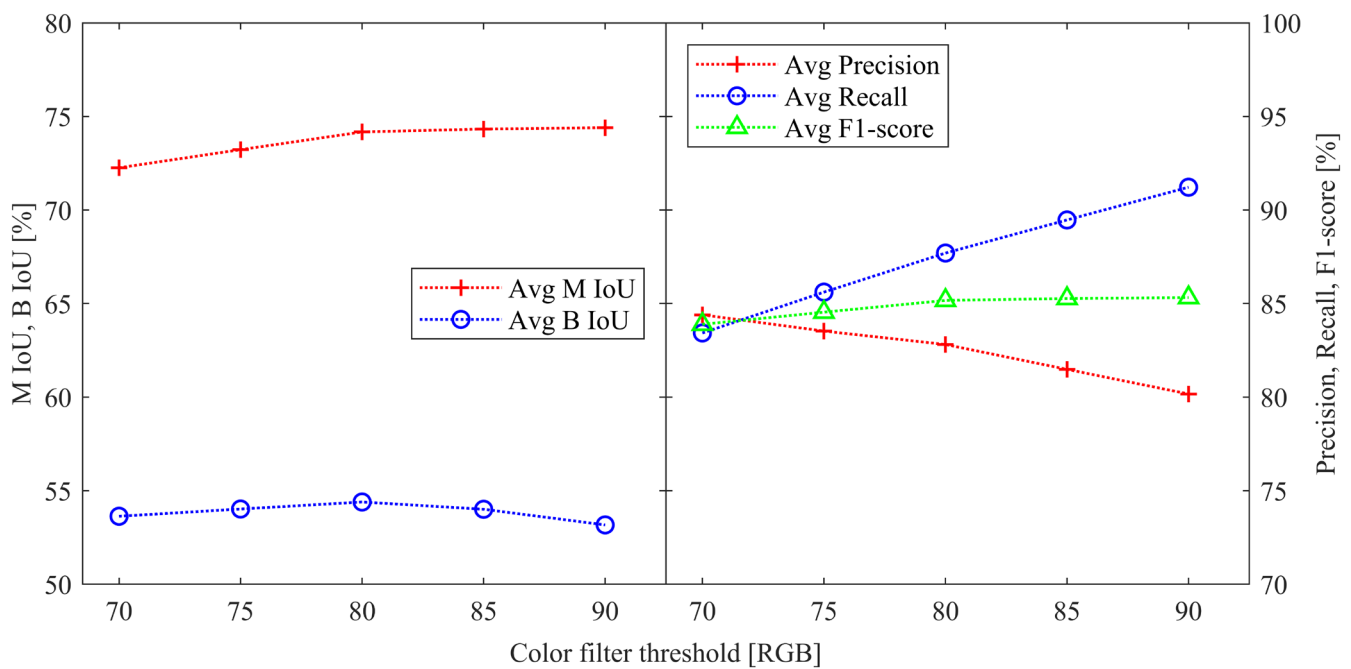


FIGURE 5 | A sensitivity analysis of the color filter conducted on the test seedling (Spruce 1).



FIGURE 6 | Generated point clouds. Comparison of the resulting dense point cloud from different photogrammetry software and 3D scanning. From left to right: Agisoft, Meshroom, RealityCapture, Shining 3D and the original seedling.

TABLE 2 | The statistics of the comparison of the different photogrammetry software and 3D scanning.

Software	Computational time (h)	Camera alignment success rate (%)	Points (10 ³)
Agisoft	22.5	100	14,440
Meshroom	5	100	1480
RealityCapture	1.5	100	5530
3D scanning	<0.3	—	10,640

Despite the large differences in initial point cloud density, the resulting mesh sizes are of similar order for both methods. The final meshes typically contain on the order of 10^5 vertices and faces, indicating that the meshing and filtering pipeline normalizes differences in input density. However, Agisoft-based reconstructions generally result in slightly larger and more detailed meshes compared to Meshroom.

In Appendix B images of the final textured meshes are shown. The final mesh of a seedling is regarded as a single object, even though it is discontinuous and composed of multiple mesh islands, none of which are solid.

In Figure 7 the evolution of the reconstruction process is visualized for the test seedling (Spruce 1)-first the unfiltered dense

TABLE 3 | Point cloud and mesh statistics throughout processing steps. Agisoft (A) and Meshroom (M) were used for photogrammetry, CloudCompare for point cloud (pcl) filtering where step one means after manual cropping and step two means after color filtering. The final meshes were created in Blender.

Photogrammetry points (10 ³)		Pcl filtering step one/ two (points, 10 ³)	Final mesh vertices/ edges/faces (10 ³)
Pine 1 A	23,030	8526/2818	357/981/617
M	4730	781/400	269/791/519
Pine 2 A	21,594	5036/2480	306/830/517
M	4708	624/373	243/716/469
Pine 3A	14,641	3558/1322	197/539/336
M	4601	401/218	152/449/294
Spruce 1 A	14,436	2948/1344	220/650/441
M	1639	429/227	168/529/360
Spruce 2 A	18,067	5047/2346	376/1110/748
	4643	634/376	273/865/587
Spruce 3 A	15,029	4375/2120	326/959/645
M	4598	613/355	252/797/542

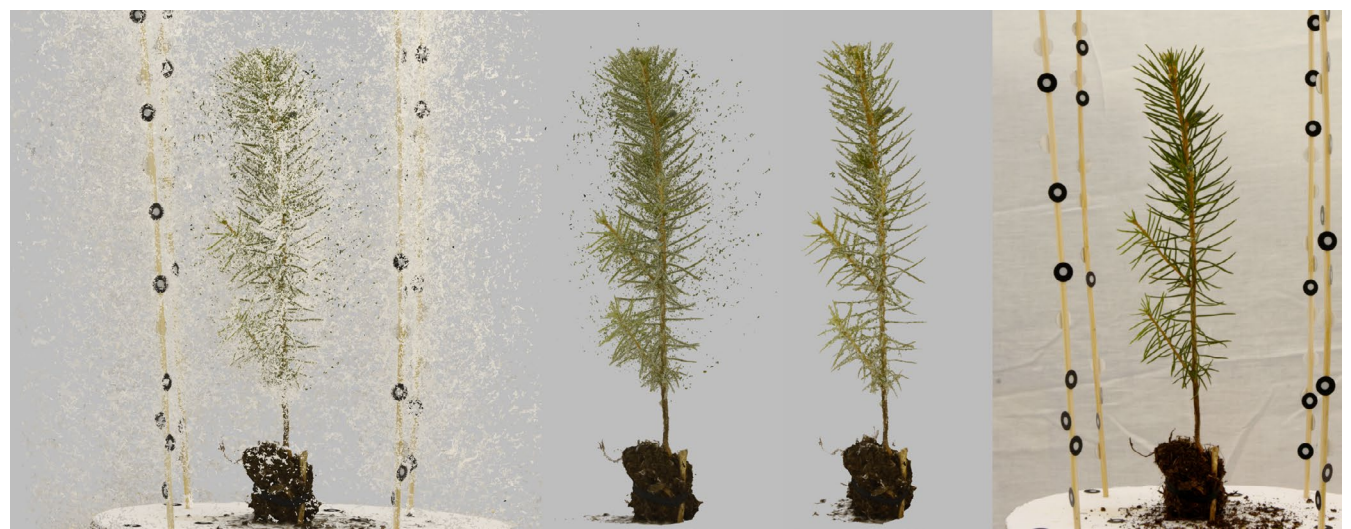


FIGURE 7 | Evolution of reconstruction. The evolution from unfiltered dense point cloud generated in Agisoft, to manually cropped and color filtered point cloud in CloudCompare, to finished mesh in Blender, and finally the original seedling (spruce 1) captured by the DSLR-camera (in order from left to right).

point cloud—resulting from the photogrammetry software (Agisoft), then the point cloud after filtering steps and finally the mesh compared to a photo of the original seedling.

3.1.1 | Quality Assessment: Agisoft

The mask IoU (M IoU) and boundary IoU (B IoU) for each seedling are presented in Figure 8. The results indicate similar performance between spruce and pine, with inter-species differences being smaller than intra-species variability. The average M IoU is 75.6% for spruce and 75.8% for pine, while the corresponding B IoU values are 50.3% and 49.4%, respectively. The mask- and boundary IoU across all samples averaged to 75.7% and 49.9%.

In Figure 9, the precision, recall, and F1-score for each seedling are presented. The results indicate a trade-off between precision and recall across species. Spruce exhibits higher precision, with an average of 85.5%, compared to 81.4% for pine. In contrast, pine achieves higher recall, with an average of 91.7% compared to 86.9% for spruce. Despite these differences, the F1-score is identical for both species, with an average of 86.1%, indicating balanced overall performance. The precision and recall for all samples averaged to 83.5% and 89.3%, respectively.

The higher recall for pine suggests improved detection completeness, while the higher precision for spruce indicates fewer false positives. The similar F1-scores imply that the segmentation method performs consistently across species, with differences

primarily related to the precision–recall trade-off rather than overall accuracy.

3.1.2 | Quality Assessment: Meshroom

The mask IoU (M IoU) and boundary IoU (B IoU) for each seedling are presented in Figure 10. The results indicate similar

performance between spruce and pine, with inter-species differences being smaller than intra-species variability. The average M IoU is 69.9% for spruce and 71.1% for pine, while the corresponding B IoU values are 42.7% and 40.2%, respectively. The mask- and boundary IoU across all samples averaged to 70.5% and 41.5%.

In Figure 11, the precision, recall, and F1-score for each seedling are presented. As for Agisoft, a trade-off between

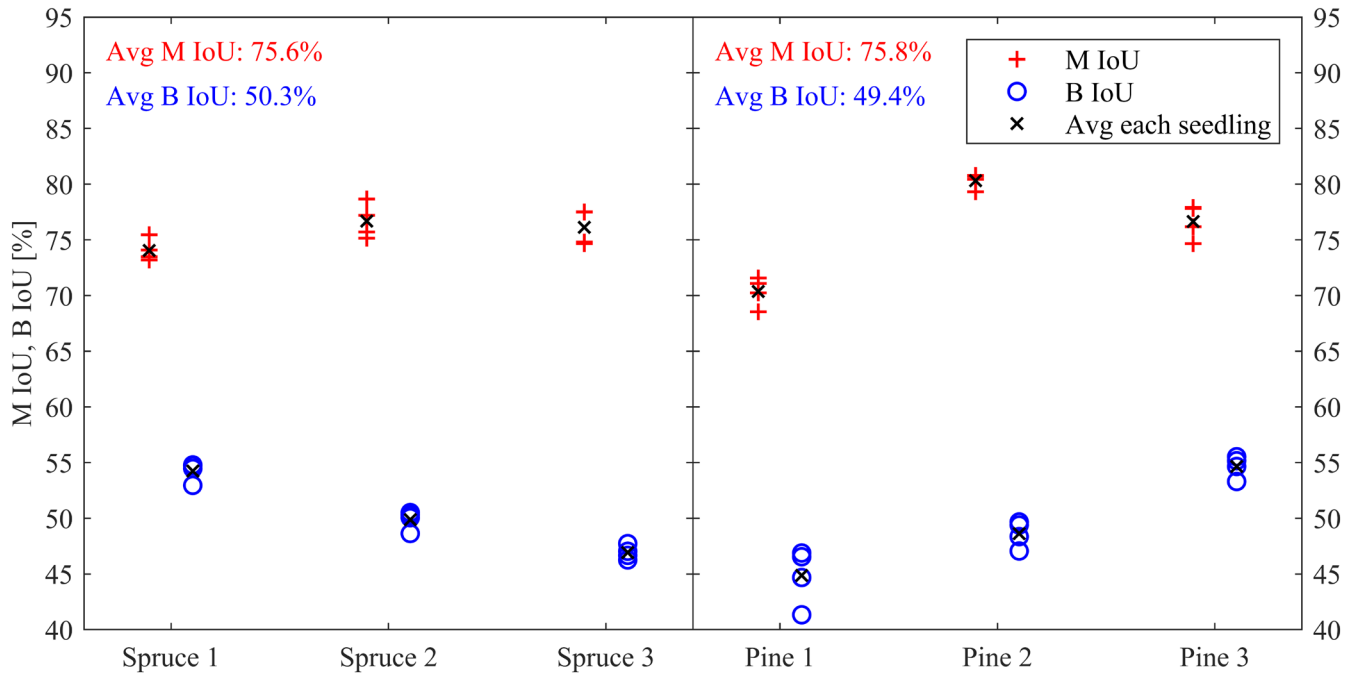


FIGURE 8 | Mask- and boundary IoU. Mask IoU (M IoU, red plus signs) and boundary IoU (B IoU, blue circles) values presented for each seedling, with spruce gathered to the left graph and pine to the right graph. The black crosses show the average across the four viewing angles for each seedling and metric.

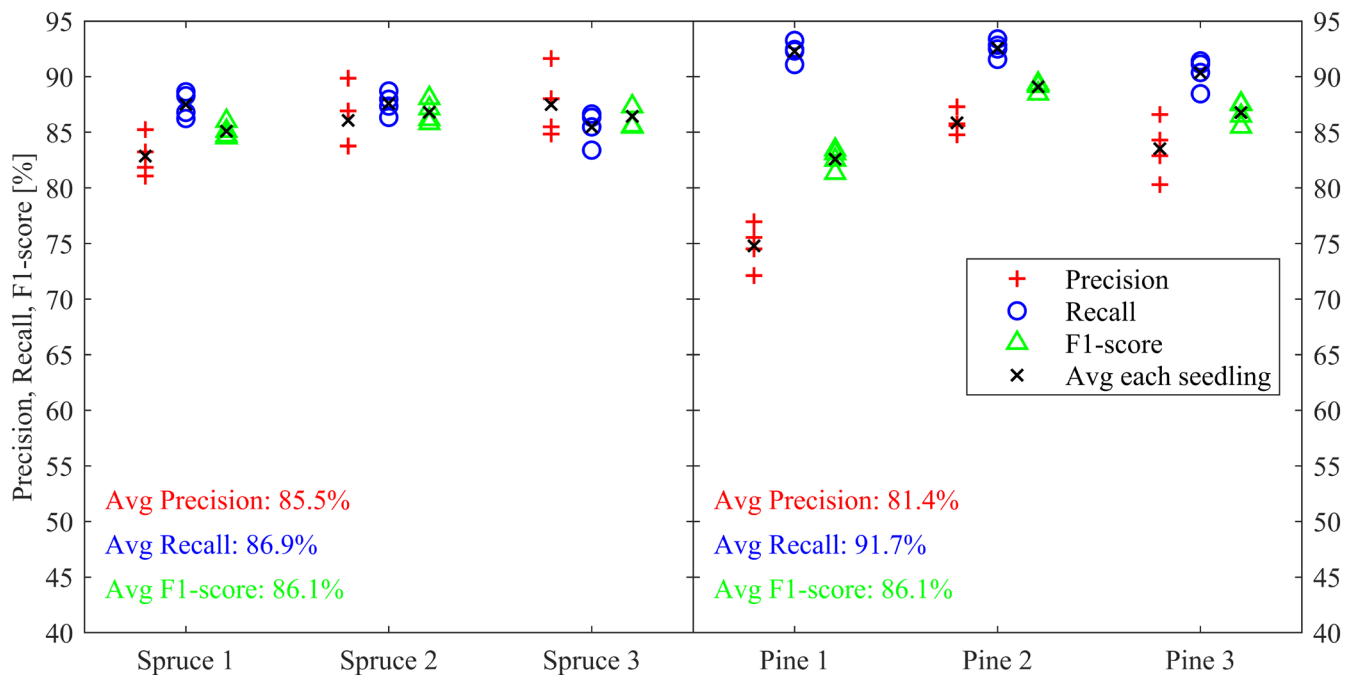


FIGURE 9 | Precision, recall and F_1 -score. Precision (red plus signs), recall (blue circles) and F_1 -score measured values presented for each seedling, with spruce gathered to the left graph and pine to the right graph. The black crosses show the average across the four viewing angles for each seedling and metric.

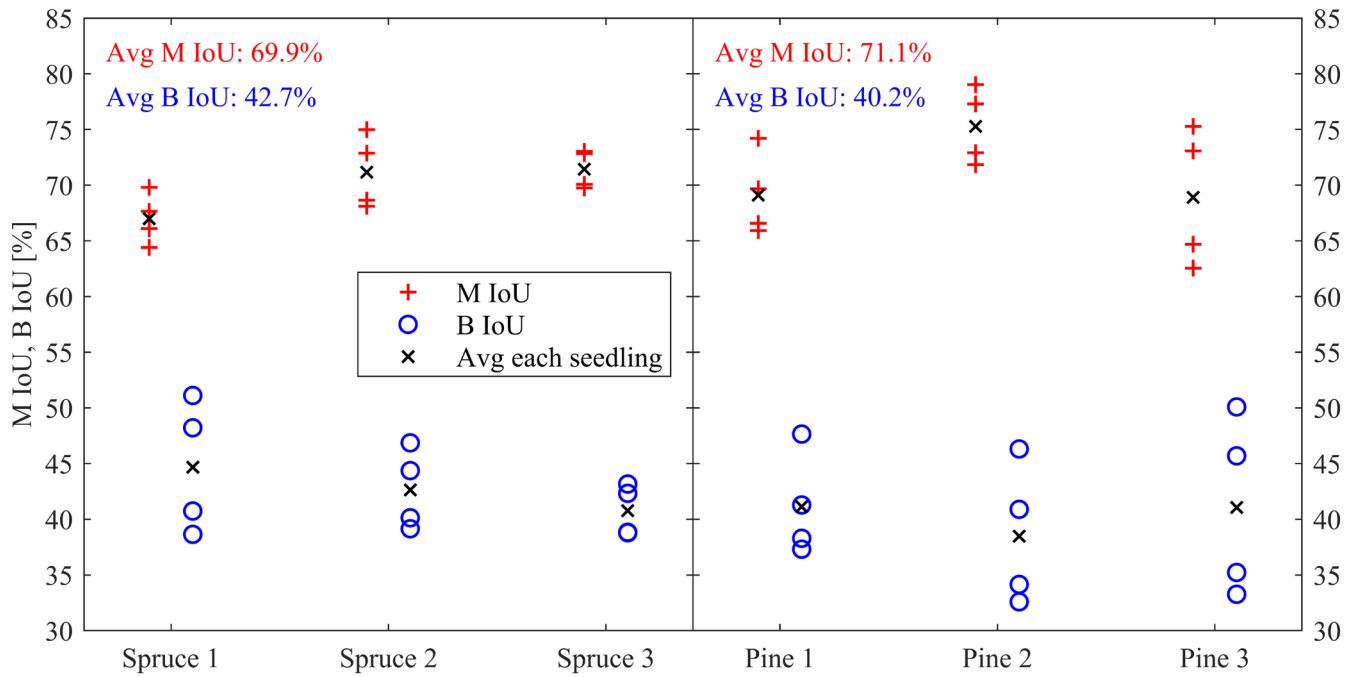


FIGURE 10 | Mask- and boundary IoU, Meshroom. Mask IoU (M IoU, red plus signs) and boundary IoU (B IoU, blue circles) values presented for each seedling, with spruce gathered to the left graph and pine to the right graph. The black crosses show the average across the four viewing angles for each seedling and metric.

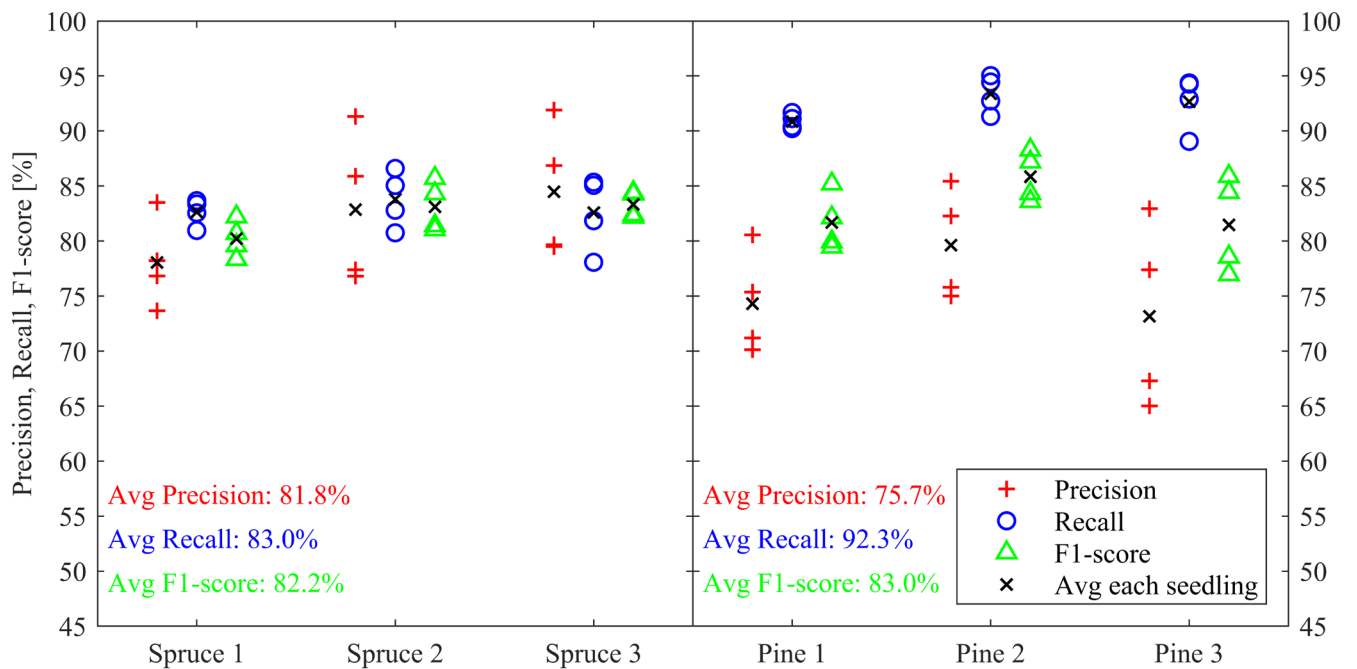


FIGURE 11 | Precision, recall and F_1 -score, Meshroom. Precision (red plus signs), recall (blue circles) and F_1 -score measured values presented for each seedling, with spruce gathered to the left graph and pine to the right graph. The black crosses show the average across the four viewing angles for each seedling and metric.

precision and recall is observed. Spruce exhibits higher precision, with an average of 81.8%, compared to 75.7% for pine. In contrast, pine achieves higher recall, with an average of 92.3% compared to 83.0% for spruce. The F_1 -score is similar between

species, with averages of 82.2% for spruce and 83.0% for pine, indicating comparable overall performance. The precision, recall, and F_1 -score for all samples averaged to 78.8%, 87.7%, and 82.6% respectively.

3.1.3 | Comparison of Agisoft Versus Meshroom

Compared to Agisoft, Meshroom yields lower performance across all metrics, with differences of -5.2 , -8.4 , -4.7 , -1.6 and -3.5 percentage points for mask IoU, boundary IoU, precision, recall, and F1-score, respectively.

Both software result in higher recall for pine, which suggests improved detection completeness, while the higher precision for spruce indicates fewer false positives. The similar F1-scores imply that the workflow performs consistently across species and software, with differences primarily related to the precision–recall trade-offs.

3.2 | Workflow for 3D Reconstruction

The finalized semi-automated workflow for digital 3D reconstruction of small, flexible objects, as developed in this work, is presented in Figure 12. The initial data acquisition step is performed via an automated, robotic DSLR camera setup which provides images to the photogrammetry software. The images are then aligned and used for generating a dense point cloud of the object. Next, the point cloud is imported to a point cloud processing software where it manually is cropped to the region of interest before a color filter is applied. The filtered point cloud is imported to a mesh processing software where three steps are applied, i.e., decimating the number of points by distance merging, turning the points into mesh, and finally, mesh cleaning. The resulting textured mesh can then be exported and saved in several formats for usage in compatible programs. The execution times mentioned in the workflow diagram are approximated maximum times it takes to perform the steps (with chosen hardware see specification in Appendix A), and both the filtering and meshing steps take less than 5 min each for an experienced user. However, the photogrammetry (reconstruction) stage takes the majority of the time, and can be reduced depending on settings and software.

4 | Discussion

This work has targeted the generation of digital seedling representations that can enable image-based computer vision models

in synthetic environments. For the intended application, both geometric fidelity and visual realism are expected to influence performance, although the relative importance of these factors remains to be investigated. In the present study, no quantitative assessment of texture fidelity has been performed, and the discussion is therefore limited primarily to geometric reconstruction quality and visual appearance. A reconstruction can therefore be considered sufficiently accurate if it captures the overall structure, silhouette, and visual appearance of the object such that it can be recognized and used for training data generation.

Difficulties were encountered when using the 3D scanner to maintain consistent tracking while keeping the correct scanning distance. This can be attributed to the small scale and slender geometry of the objects, which provide limited geometric features for stable tracking. Increasing the coverage of the predefined scanning path could have reduced the need for manual control of the robot arm, potentially improving tracking stability and thus the quality of the resulting point cloud. The 3D scanner produced significantly better results when the seedling was coated in white powder and using the laser scan mode, suggesting that the natural surface of the seedling is troublesome for laser-based scanning. This is likely attributed to the surface being too reflective. This gives a hint of the resolution needed for 3D scanning fine structures, such as spruce needles, as laser did succeed with the coating applied compared to the rapid scan, see Table 1 for comparison of scan mode resolutions and accuracy. However, we did not investigate this approach further since the laser scanning cannot capture color and texture, and the coating we had available was destructive for the seedling. The photogrammetry approach was therefore selected for further research since it produced higher-quality point clouds without altering the object, while preserving color and texture information.

All the examined photogrammetry approaches rendered point clouds with an adequate representation of the stem but lacked data for the needles especially, but also to some extent for the branches. The default settings were used for each software, during the comparison, except any filtering settings, which were turned off when possible. Although alternative parameter configurations may improve results, systematic parameter optimization was outside the scope of this study.

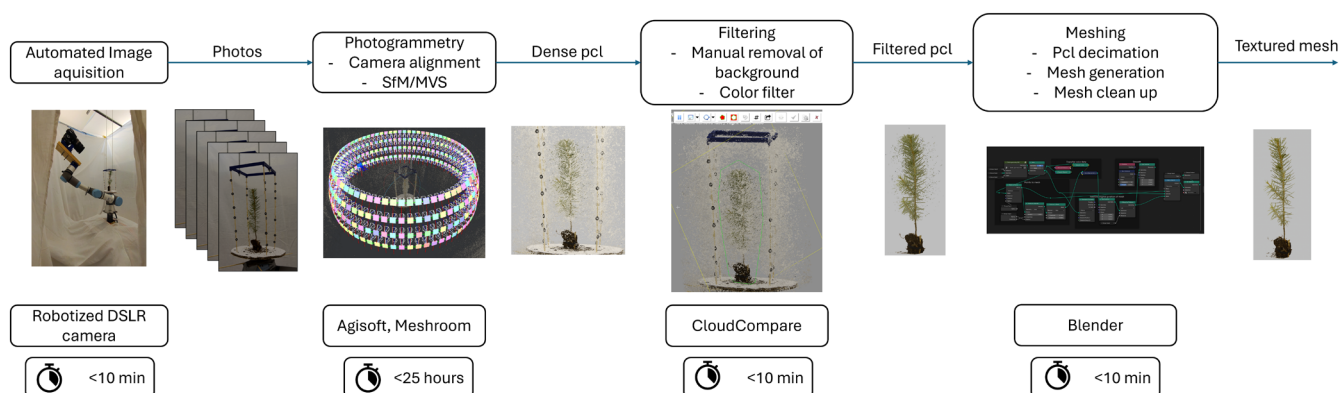


FIGURE 12 | The proposed workflow. The proposed workflow for digital reconstruction of small flexible objects. In our work, this workflow was tested specifically for conifer seedlings.

Although all the photogrammetry and 3D scanner software included mesh generation functions, the resulting meshes were generally of poor quality, capturing only stems, and in the best cases, also partial branches and needles. The same limitations were observed in the point clouds generated with filtering applied; hence, the choice was made to generate unfiltered point clouds for further processing in external software. However, filtering comes with challenges and trade-offs between accuracy and data loss. When applying the color filter, a threshold value determines whether a point is retained or discarded. In reality, a sharp object outline is blurred because of natural shades in the color and also because one pixel can represent the average color of both the object edge and background. This phenomenon is more prominent for smaller objects as the effective resolution decreases. A similar uncertainty is also present in the ground truth masks, which introduces a systematic bias in the evaluation metrics. It was observed that our choice of threshold for the color filter generally led to thinner reconstructions of the needles, but with better overall reconstruction accuracy.

During this research it became clear that the choice of camera and camera parameters are decisive for the performance of the photogrammetry software. The camera settings were chosen by iterative testing, and the trade-off on quality and sufficient lighting when adjusting iso, shutter speed and f -value were significant for the quality of the resulting point cloud. Moreover, sensor-induced artifacts led to poor camera alignment in the different software when using unprocessed RAW images and therefore the built-in camera JPEG format with precalibrated noise removal was used. Although JPEG format introduces data loss and minor color inconsistencies, it gave sufficient detailed visualization and resulted in good camera alignment in the photogrammetry software. However, the accuracy and level of detail in the resulting point cloud depend highly on both sensor and lens characteristics. These parameters therefore need to be optimized to suit the object size and photographing distance, accounting for lens distortion and required level of detail. Result improvements were also observed when altering the photographing environment by increasing the lighting and minimizing disturbing features in the background. These observations indicate that the method is sensitive to imaging conditions, although this effect has not been quantified in this study.

Several methods were tested for converting the point cloud into a mesh, before ultimately using Blender for the task. Small and slender structures are challenging to mesh, and the outcome highly depends on the quality of the point cloud. For example, when substantial noise is present between several thin structures it is impossible for the meshing algorithm to decide whether the points belong to one or multiple surfaces, or in which direction the mesh should be inferred. The difficulty is amplified when needles intersect or touch each other, resulting in a compound mesh of several needles when they actually should be separated. That effect, in turn, complicates subsequent processing steps of mesh flattening, spike removal, etc. Another example is when data is missing from the point cloud, which often results in separate mesh islands. Creating a single continuous mesh by closing these gaps can be a heavy or complex task and has not been implemented in this workflow, since that is not crucial for the intended use of the textured mesh in a virtual environment. Nevertheless, the developed Blender workflow produced stable

results for conifer seedlings under consistent settings. When switching between spruce and pine seedlings, adjusting only two values further enhanced stability. This indicates that the workflow may be applicable to other small, slender structured objects by modifying just these two parameters: mesh voxel size and radius.

The results indicate comparable reconstruction performance between pine and spruce seedlings for both Agisoft and Meshroom, with inter-species differences generally smaller than the variability observed within each species. No consistent advantage for either species is observed in terms of mask IoU, boundary IoU, or F1-score, suggesting that the reconstruction pipeline is robust across both plant types. For both photogrammetry approaches, a systematic trade-off between precision and recall is observed. Spruce exhibits higher precision, indicating fewer false positives, whereas pine achieves higher recall, reflecting a more complete reconstruction of the seedling geometry. Despite this, the F1-scores are similar between species, indicating that these differences largely balance out in terms of overall performance. Comparing the two pipelines, Agisoft consistently yields higher performance than Meshroom across all metrics, particularly for mask IoU and boundary IoU. The largest differences are observed for boundary IoU, indicating that Agisoft more accurately captures fine-scale geometric details. Meshroom, while producing lower IoU values, maintains a similar precision–recall balance, suggesting that the primary differences between the methods lie in geometric fidelity and point cloud density. The observed precision–recall trade-off can be partially explained by structural differences between the species. Pine seedlings, with longer needles and a more open structure, provide more distinct features across views, which may facilitate feature matching and improve reconstruction completeness (recall). In contrast, spruce seedlings exhibit denser foliage and increased self-occlusion, which can limit visibility of certain regions but may reduce ambiguity in detected structures, contributing to higher precision. However, the relatively small magnitude of these differences suggests that biological factors play a secondary role compared to the characteristics of the reconstruction pipeline. Overall, the boundary IoU was lower than the mask IoU, indicating that the reconstructions did not accurately follow the true outlines. A plausible explanation is the effect of the color filter, which tends to thin the needles. This has a pronounced influence on the boundary IoU but only a limited effect on the mask IoU.

The metrics reveal that the reconstructions are not perfect replicas of the true objects; however, that was never the main goal of this work. Instead, the aim was to produce digital versions that are sufficiently representative and that, in future work, could be used to create datasets for deep learning models. Through qualitative assessment of the reconstructed meshes, it is possible to distinguish the species and confirm that each mesh corresponds to a seedling. For the intended application, both geometric fidelity and visual appearance are relevant, and aspects such as accurate silhouettes and preserved texture are expected to play an important role. At the same time, the impact of fine-scale geometric details, such as precise needle thickness or small-scale topology, remains unclear and has not been explicitly evaluated in this study.

It is difficult to compare the reconstruction results to related research since the quantitative metrics differ a lot between works. However, the proposed method of evaluating the reconstruction quality could improve the comparability of related works. It produces a stable and reliable ground truth compared to, e.g., Hu et al. (2024) that relies on RealityCapture to generate a ground truth and (Yang et al. 2024) that does not compare the point cloud quality against a ground truth at all. Note that the evaluation method is not part of the presented semi-automated reconstruction workflow. Nevertheless, the evaluation is limited to image-space metrics and does not capture full 3D geometric accuracy due to the lack of reliable 3D ground truth intrinsic for flexible objects. Furthermore, no dedicated evaluation of texture or color fidelity has been conducted other than qualitative comparisons, which also is a noted drawback in other works (Wang et al. 2024; Rajapaksha 2025). It would be of interest to compare the proposed workflow with recent deep learning-based approaches such as Neural Radiance Fields (NeRF). These methods can generate photorealistic novel views and reconstructions that are competitive with traditional photogrammetry, particularly for complex plant structures (Hu et al. 2024). However, several limitations remain, including high computational cost, sensitivity to imaging conditions, and challenges in reconstructing accurate geometry in complex scenes or under insufficient sampling (Rajapaksha 2025). Furthermore, the implicit volumetric nature of NeRFs creates polygonal meshing incompatibility issues that need to be solved for its use in virtual environments (Frasiolas et al. 2025). In contrast, photogrammetry remains a mature and well-established method capable of producing dense and geometrically consistent reconstructions, particularly when sufficient surface texture is available, and benefits from robust, reproducible workflows across a wide range of applications (Remondino et al. 2023).

A significant advantage of the presented methodology is that keeping the scanned object stationary while moving the camera around it reduces the risk of motion-induced errors, discussed in (Yang et al. 2024), and preserves the structural integrity of delicate or flexible objects. This approach also ensures consistent lighting conditions and stable reference points for improved reconstruction accuracy. Compared to the reconstruction of conifer shoots in (Pisek et al. 2023), the presented method has the advantage of producing the results without manual processing and without adjusting the scanning parameters, while also retaining texture and color. In contrast to prior approaches that require manual intervention or do not preserve visual information, the proposed workflow enables a semi-automated reconstruction process suitable for dataset generation. As shown, the methodology is effective for the digital reconstruction of small and slender structured flexible objects and holds potential for increased automation. While most steps are already automated or can be automated with existing software functionality, the manual transfer of files between applications remains the primary bottleneck that needs to be addressed to achieve a fully automated workflow.

5 | Conclusions

The accuracy and quality of point clouds generated through photogrammetry is highly dependent on the quality of the image

capturing device. However, for the purpose of generating auto annotated images to develop detectors for computer vision, a decent hobby level DSLR camera showed feasible to reconstruct six samples of conifer saplings with two of three tested photogrammetry software. The 3D scanner approach explored did not produce satisfactory results.

This study presents a semi-automated workflow for 3D reconstruction of small and slender structured flexible objects that are sensitive to movement and vibrations. The workflow combines automatic image acquisition, photogrammetric reconstruction, filtering and meshing to produce textured meshes. Based on our best results from this reconstruction and evaluation methodology six conifer seedlings were digitally reconstructed with an average mask IoU-value of 75.7% and an average F1-score of 86.1%. The sample size is limited, and further experiments, including additional conifer seedlings and other object types, are required to validate the workflow.

Acknowledgements

Declaration of generative AI and AI-assisted technologies in the writing process: During the preparation of this work the author(s) used ChatGPT-5 in order to translate and improve the language and structure of the written text. After using this tool/service, the author(s) reviewed and edited the content as needed and take full responsibility for the content of the publication.

Funding

This work was supported by the Interreg Aurora under Grant NYPS 20357984; Jubileumsfonden at Luleå University of Technology under Grant JCSMKJF23-0004; and VINNOVA under Grant 2023-02747.

Ethics Statement

This study did not involve human participants or animals. The plant material used does not fall under protected or regulated species, and no specific permits or ethical approvals were required. Therefore, no ethical approval or informed consent was required.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Raw image data and finalized textured meshes are freely available at Zenodo.org with <https://doi.org/10.5281/zenodo.19823955>.

References

- Abbood, S. A., H. A. Ajjah, A. H. H. Alboabidallah, M. U. Mohammed, and F. M. Alattar. 2024. "Artificial Intelligence Techniques in Photogrammetry Application: A Review." In *Presented at the AIP Conference Proceedings*. American Institute of Physics. <https://doi.org/10.1063/5.0212428>.
- Blender Documentation Team. n.d. "Blender 4.2 Manual." [Blender.org](https://docs.blender.org/manual/en/4.2/index.html). Accessed April 14, 2026. [Online]. <https://docs.blender.org/manual/en/4.2/index.html>.
- Cheng, B., R. Girshick, P. Dollár, A. C. Berg, and A. Kirillov. 2021. "Boundary IoU: Improving Object-Centric Image Segmentation Evaluation." In *Presented at the Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*,

- 15329–15337. IEEE Computer Society. <https://doi.org/10.1109/CVPR46437.2021.01508>.
- Cova, G. R., S. J. Prichard, E. Rowell, et al. 2023. “Evaluating Close-Range Photogrammetry for 3D Understory Fuel Characterization and Biomass Prediction in Pine Forests.” *Remote Sensing* 15, no. 19: 4837. <https://doi.org/10.3390/rs15194837>.
- De Paolis, L. T., V. De Luca, C. Gatto, G. D’Errico, and G. I. Paladini. 2020. “Photogrammetric 3d Reconstruction of Small Objects for a Real-Time Fruition.” In *Presented at the Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 375–394. Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-030-58465-8_28.
- Dosovitskiy, A., G. Ros, F. Codevilla, A. Lopez, and V. Koltun. 2017. “CARLA: An Open Urban Driving Simulator.” In *Conference on Robot Learning*, edited by S. Levine, V. Vanhoucke, and K. Goldberg, Proceedings of Machine Learning Research, vol. 78. Jmlr-Journal Machine Learning Research.
- Frasiolas, P., G.-A. Cheimariotis, P. K. Papadopoulos, and D. Zarpalas. 2025. “HAME-NeRF: High Accuracy Mesh Extraction Leveraging Neural Radiance Fields.” In *Presented at the International Conference on Computer Analysis of Images and Patterns*, 330–340. Springer. <https://doi.org/10.1109/AIXVR63409.2025.00024>.
- Galantucci, L. M., M. G. Guerra, and F. Lavecchia. 2018. “Photogrammetry Applied to Small and Micro Scaled Objects: A Review.” In *Lecture Notes in Mechanical Engineering*, 57–77. Pleiades Publishing. https://doi.org/10.1007/978-3-319-89563-5_4.
- Griwodz, C., S. Gasparini, L. Calvet, et al. 2021. “AliceVision Meshroom. Proceedings of the 12th ACM Multimedia Systems Conference.” 241–247. <https://doi.org/10.1145/3458305.3478443>.
- Ham, H., and J. Wesley. 2019. “Computer Vision Based 3D Reconstruction: A Review.” *International Journal of Electrical and Computer Engineering (IJECE)* 9, no. 4: 2394–2402. <https://doi.org/10.11591/ijece.v9i4.pp2394-2402>.
- Hansson, L. J., G. Sten, M. Rossander, et al. 2024. “Autoplant—Autonomous Site Preparation and Tree Planting for a Sustainable Bioeconomy.” *Forests* 15, no. 2: 263. <https://doi.org/10.3390/f15020263>.
- Hartley, R., and A. Zisserman. 2003. *Multiple View Geometry in Computer Vision*. Cambridge university press.
- Hu, K., W. Ying, Y. Pan, H. Kang, and C. Chen. 2024. “High-Fidelity 3D Reconstruction of Plants Using Neural Radiance Fields.” *Computers and Electronics in Agriculture* 220: 108848. <https://doi.org/10.1016/j.compag.2024.108848>.
- Iglhaut, J., C. Cabo, S. Puliti, L. Piermattei, J. O’Connor, and J. Rosette. 2019. “Structure From Motion Photogrammetry in Forestry: A Review.” *Current Forestry Reports* 5, no. 3: 155–168. <https://doi.org/10.1007/s40725-019-00094-3>.
- James, N., A. Adkinson, and A. Mast. 2023. “Rapid Imaging in the Field Followed by Photogrammetry Digitally Captures the Otherwise Lost Dimensions of Plant Specimens.” *Applications in Plant Sciences* 11, no. 5: e11547. <https://doi.org/10.1002/aps3.11547>.
- Karty, J., and B. Hament. 2024. “3D Data Collection for Individual Plant Farming.” In *2024 Systems and Information Engineering Design Symposium, SIEDS 2024*, 292–295. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/SIEDS61124.2024.10534658>.
- Krizhevsky, A., I. Sutskever, and G. E. Hinton. 2017. “ImageNet Classification With Deep Convolutional Neural Networks.” *Communications of the ACM* 60, no. 6: 84–90. <https://doi.org/10.1145/3065386>.
- LeCun, Y., Y. Bengio, and G. Hinton. 2015. “Deep Learning.” *Nature* 521, no. 7553: 436–444. <https://doi.org/10.1038/nature14539>.
- Lehto, M., H. Lideskog, and M. Karlberg. 2026. “Log Detection for Autonomous Forwarding Using Auto-Annotated Data From a Real-Time Virtual Environment.” *Journal of Terramechanics* 121: 101096. <https://doi.org/10.1016/j.jterra.2025.101096>.
- Luhmann, T., S. Robson, S. Kyle, and J. Boehm. 2014. *Close-Range Photogrammetry and 3d Imaging*. 2nd ed. De Gruyter textbook. Walter De Gruyter GmbH.
- Ma, Y., S. Soatto, J. Kosecka, and S. S. Sastry. 2012. *An Invitation to 3-d Vision: From Images to Geometric Models*. Vol. 26. Springer Science & Business Media.
- Magnuson, R., Y. Erfanifard, M. Kulicki, et al. 2024. “Mobile Devices in Forest Mensuration: A Review of Technologies and Methods in Single Tree Measurements.” *Remote Sensing* 16, no. 19: 3570. <https://doi.org/10.3390/rs16193570>.
- Marks, E., F. Magistri, and C. Stachniss. 2022. “Precise 3D Reconstruction of Plants From UAV Imagery Combining Bundle Adjustment and Template Matching.” In *2022 International Conference on Robotics and Automation (ICRA)*, 2259–2265. IEEE. <https://doi.org/10.1109/ICRA46639.2022.9811358>.
- Medina, J. J., J. M. Maley, S. Sannapareddy, N. N. Medina, C. M. Gilman, and J. E. McCormack. 2020. “A Rapid and Cost-Effective Pipeline for Digitization of Museum Specimens With 3D Photogrammetry.” *PLoS One* 15, no. 8 August: e0236417. <https://doi.org/10.1371/journal.pone.0236417>.
- Mortezapoor, S., C. Schonauer, J. Ruggeberg, and H. Kaufmann. 2022. “Photogramobot: An Autonomous ROS-Based Mobile Photography Robot for Precise 3D Reconstruction and Mapping of Large Indoor Spaces for Mixed Reality.” In *Presented at the Proceedings—2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops, VRW 2022*, 101–107. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/VRW55335.2022.00033>.
- Muthu, S., T. Wedutenko, J. Tong, C. Nguyen, and L. Petersson. 2023. “Towards End-to-End Automatic Insect Handling and Insect Scanning.” In *Presented at the 2023 International Conference on Digital Image Computing: Techniques and Applications, DICTA 2023*, 168–175. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/DICTA60407.2023.00031>.
- Okura, F. 2022. “3D Modeling and Reconstruction of Plants and Trees: A Cross-Cutting Review Across Computer Graphics, Vision, and Plant Phenotyping.” *Breeding Science* 72, no. 1: 31–47. <https://doi.org/10.1270/jsbbs.21074>.
- Over, J.-S. R., A. C. Ritchie, C. J. Kranenburg, et al. 2021. “Processing Coastal Imagery With Agisoft Metashape Professional Edition, Version 1.6—Structure from Motion Workflow Documentation. Open-File Report.” <https://doi.org/10.3133/ofr20211039>.
- Pisek, J., O. Borysenko, R. Janoutová, and L. Homolová. 2023. “Estimation of Coniferous Shoot Structure by High Precision Blue Light 3D Photogrammetry Scanning.” *Remote Sensing of Environment* 291: 113568. <https://doi.org/10.1016/j.rse.2023.113568>.
- Plum, F., and D. Labonte. 2021. “scAnt—An Open-Source Platform for the Creation of 3D Models of Arthropods (And Other Small Objects).” *PeerJ* 9: e11155. <https://doi.org/10.7717/peerj.11155>.
- Rajapaksha, U. P. 2025. “Image-based 3D Reconstruction of Complex Plant Structures.” <https://doi.org/10.60867/00000077>.
- Remondino, F., A. Karami, Z. Yan, G. Mazzacca, S. Rigon, and R. Qin. 2023. “A Critical Analysis of NeRF-Based 3D Reconstruction.” *Remote Sensing* 15, no. 14: 3585. <https://doi.org/10.3390/rs15143585>.
- Remondino, F., E. Nocerino, I. Toschi, and F. Menna. 2017. “A Critical Review of Automated Photogrammetric Processing of Large Datasets.” In *Presented at the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences—ISPRS Archives*, 591–599. International Society for Photogrammetry and Remote Sensing. <https://doi.org/10.5194/isprs-archives-XLII-2-W5-591-2017>.
- Rice, J. R., G. A. Fricker, and J. Ventura. 2025. “An End-To-End Deep Learning Solution for Automated LiDAR Tree Detection in the Urban

Environment.” *ISPRS Open Journal of Photogrammetry and Remote Sensing* 17: 100092. <https://doi.org/10.1016/j.ophoto.2025.100092>.

Róg, M., and A. Rzonca. 2021. “The Impact of Photo Overlap, the Number of Control Points and the Method of Camera Calibration on the Accuracy of 3D Model Reconstruction.” *GaEE* 15, no. 2: 67–87. <https://doi.org/10.7494/geom.2021.15.2.67>.

Schmidt, H., K. Büttner, and G. Prokop. 2023. “Methods for Virtual Validation of Automotive Powertrain Systems in Terms of Vehicle Drivability—A Systematic Literature Review.” *IEEE Access* 11: 27043–27065. <https://doi.org/10.1109/ACCESS.2023.3257106>.

Shamblin, Z., and P. Puppo. 2024. “Developing a 3D Digitization Protocol at the Marshall University Herbarium Using Free, Open-Source 3D Reconstruction Software.” *Castanea* 88, no. 2: 191–204. <https://doi.org/10.2179/0008-7475.88.2.191>.

Sirotnin, D., and M. Polyakov. 2022. “Hardware and Software Development of the System for Photogrammetry of Small Objects.” In *Presented at the Proceedings—2022 2nd International Conference on Technology Enhanced Learning in Higher Education, TELE 2022*, 78–82. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/TELE55498.2022.9800950>.

Stausberg, L., B. Jost, L. Klingbeil, and H. Kuhlmann. 2024. “A 3D Surface Reconstruction Pipeline for Plant Phenotyping.” *Remote Sensing* 16, no. 24: 4720. <https://doi.org/10.3390/rs16244720>.

Wang, C., L. Tinsley, and B. H. S. Asli. 2024. “Development of a Virtual Environment for Rapid Generation of Synthetic Training Images for Artificial Intelligence Object Recognition.” *Electronics* 13: 4740. <https://doi.org/10.3390/ELECTRONICS13234740>.

Yang, D., H. Yang, D. Liu, and X. Wang. 2024. “Research on Automatic 3D Reconstruction of Plant Phenotype Based on Multi-View Images.” *Computers and Electronics in Agriculture* 220: 108866. <https://doi.org/10.1016/j.compag.2024.108866>.

Yang, T., J. Ye, S. Zhou, A. Xu, and J. Yin. 2022. “3D Reconstruction Method for Tree Seedlings Based on Point Cloud Self-Registration.” *Computers and Electronics in Agriculture* 200: 107210. <https://doi.org/10.1016/j.compag.2022.107210>.

Yin, H., S. Luo, J. Tang, and J. Hua. 2024. “RGB Camera-Based Monocular Stereo Vision Applied in Plant Phenotype: A Survey.” *Computers and Electronics in Agriculture* 227: 109523. <https://doi.org/10.1016/j.compag.2024.109523>.

Appendix A

Photogrammetry software running parameters

TABLE A.1 | The running parameters used for the seedling reconstructions in Agisoft. The parameters changed from default are accuracy, quality, and filtering mode.

<i>Point Cloud alignment parameters</i>	
Accuracy	Highest
Generic preselection	Yes
Key point limit	140,000
Tie point limit	40,000
Adaptive camera model fitting	No
<i>Dense cloud and depth map generation parameters</i>	
Quality	High
Filtering mode	Disabled
Software version	1.5.5.9097

TABLE A.2 | The running parameters for the seedling reconstruction in Meshroom. The parameters changed from default are the ones shown in the table.

<i>FeatureExtraction</i>	
Describer density and quality	High
Force CPU Extraction	No
StructureFromMotion	
Initial pair A/B	Initial pair manually chosen in case of failed alignment
<i>Meshing</i>	
Colorize output	Yes
Save Raw Dense Point Cloud	Yes
Software version	2023.3.0

TABLE A.3 | The computers used in our work. The Asus is the primary computer and the Dell is only used for running Agisoft.

Computer	Asus ROG flow x16	Dell Precision Tower 5810
CPU	13th gen Intel Core i9-13900H 14 cores, 20 threads, 24 MB Cache, 2.60 GHz	Intel Xeon E5-1603 v3 4 cores, 4 threads, 10 MB cache, 2.80 GHz
RAM	32 GB	8 GB
GPU	Nvidia Gforce RTX 4070 Laptop GPU	2 * Nvidia Quadro M2000
Clock freq. (MHz)	2175	2 * 1162
GPU Memory	8 GB	2 * 4 GB
CUDA Compute Capability per GPU	8.9	5.2

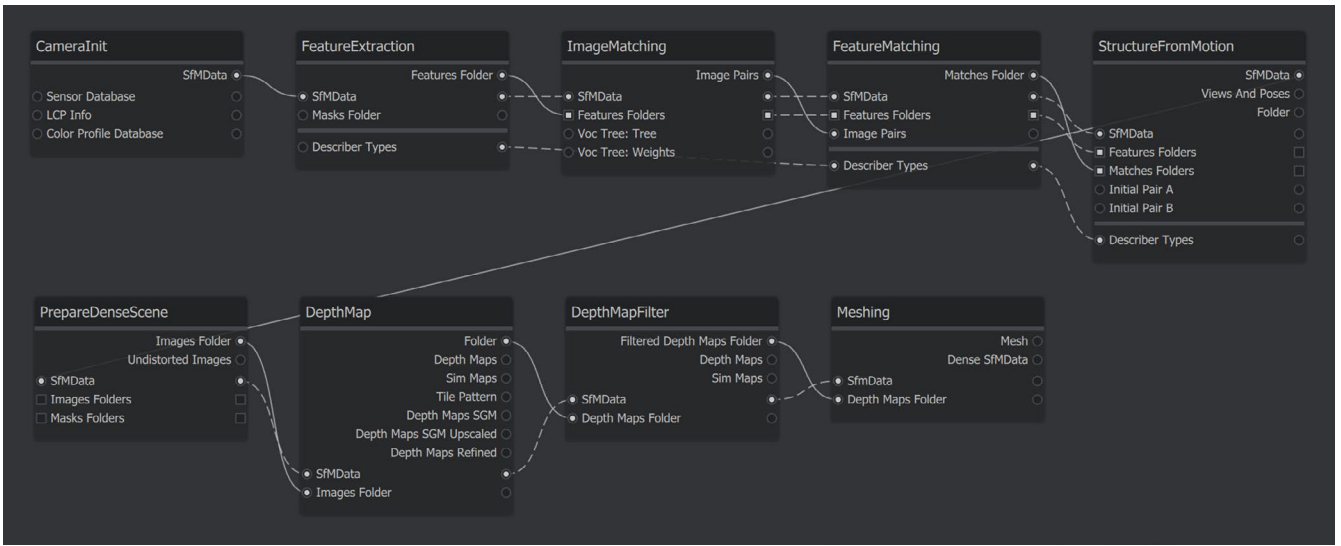


FIGURE A.1 | The pipeline used in Meshroom. Compared to the default pipeline the last two nodes, Mesh filtering and Texturing are removed since we export the results generated in the Meshing node.

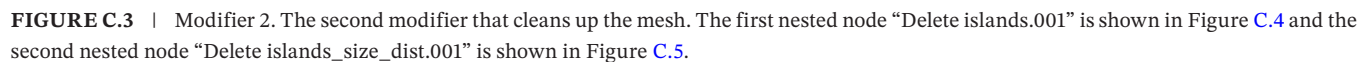
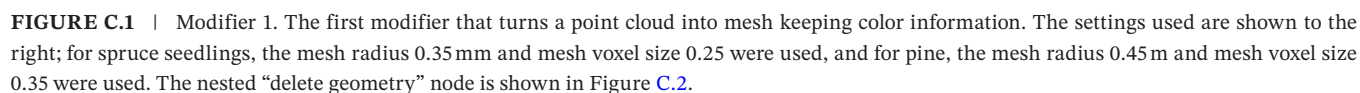
Appendix B

Comparison images of the real seedlings versus the meshes of the seedlings.

B.Appendix: Comparison images of the real seedlings vs. the meshes of the seedlings



FIGURE B.1 | Real versus meshed seedlings. Image triples of the real seedlings compared to the corresponding reconstructed digital seedlings (textured meshes). The real seedlings are in the middle of their corresponding digital versions with Agisoft to the left and Meshroom to the right. All seedlings have the same scale, relative to Spruce 1 that approximately has a height of 24 cm.



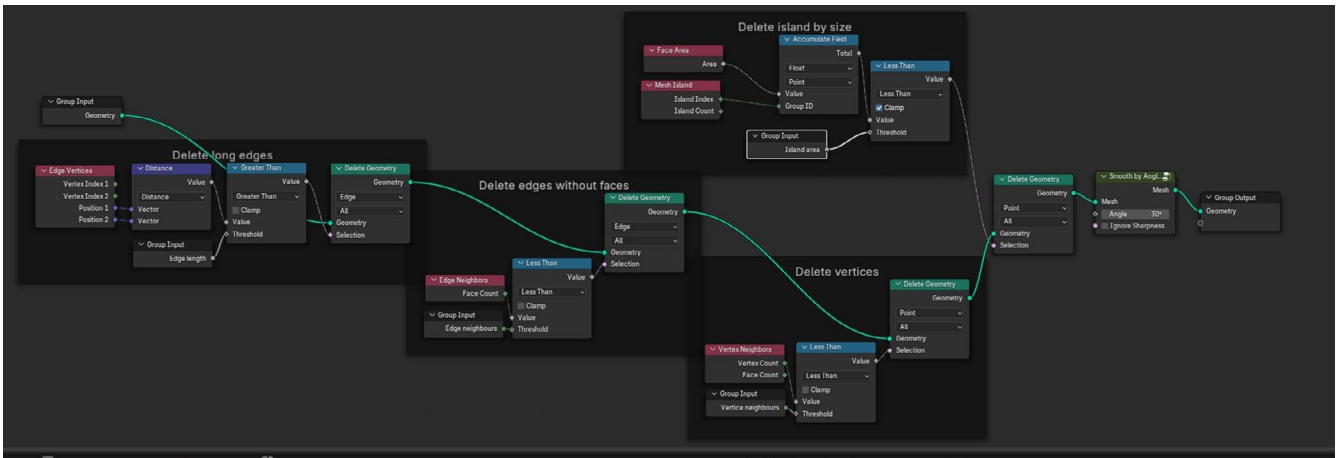


FIGURE C.4 | Modifier 2.1. The nested “Delete islands.001” node is shown here. Edges that are too long or do not have neighboring faces are deleted, as well as vertices that have too few neighboring vertices. Mesh islands that are too small are also deleted. The thresholds are shown in Figure C.3.

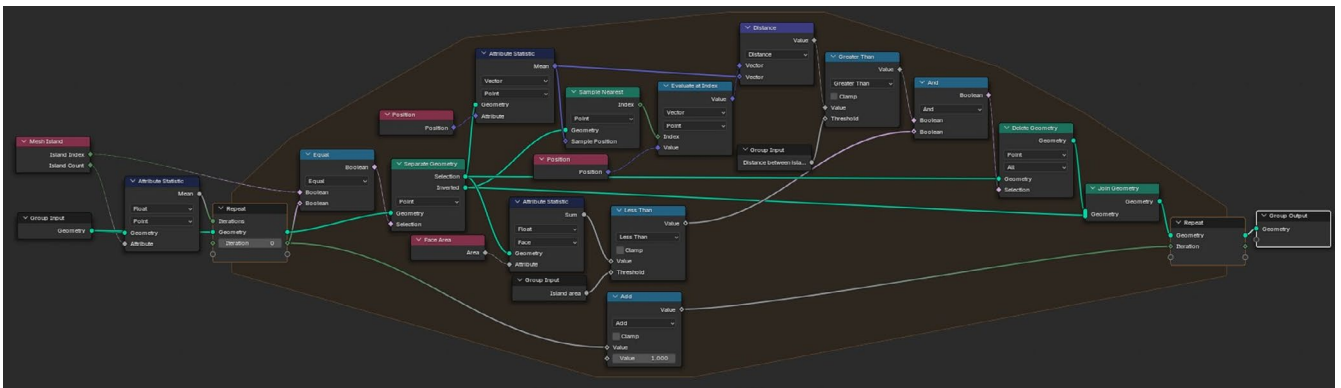


FIGURE C.5 | Modifier 2.2. The nested “Delete islands_size_dist.001” node is shown here. Mesh islands that are small and too far away from nearest neighbor are deleted. The thresholds are shown in Figure C.3.