DOI 10.5162/SMSI2025/A3.2

Application of NeRF and RGB-D Camera for the Generation of Digital Twins of Biological Soft Tissues

Marton C. Mezei[1], *Fabrizia Auletta*[1], *Alessandra Gavazza*[2], *Giacomo Rossi*[2] *and Calogero M. Oddo*[1]
[1] BioRobotics Institute, Sant'Anna School of Advanced Studies, Pisa, Italy
[2] School of Biosciences and Veterinary Medicine, University of Camerino, Matelica, Italy
Calogero.Oddo@santannapisa.it, Fabrizia.Auletta@santannapisa.it

Summary: We present an integrated scanning system designed for data collection and 3D reconstruction of macroscopic ex-vivo soft tissues. It combines a commercial RGB-D camera, a rotary stage, and a lighting system to capture high-resolution images with known and repeatable camera poses. The captured data enable the training of Neural Radiance Field (NeRF) models. The technical validation assesses the precision of the camera system and the accuracy of the NeRF-based 3D reconstruction. The demonstration of the system in clinical settings is illustrated on ex-vivo tissue scans, proving the system's potential for real-world deployment in medical environments.

Keywords: Medical Sensors, Digital Twin, Image-Based 3D Reconstruction, AI-Driven 3D Reconstruction, Data Fusion in Sensor Systems, AI in Measurement Systems, Ex-vivo Tissue Scanning.

Background, Motivation and Objective

The creation of high-fidelity digital twins and 3D virtual reconstructions of physical samples, especially in the medical field, is a rapidly advancing field. However, generating these detailed digital twins presents challenges, particularly for surfaces that are reflective, homogeneous or nearly symmetrical, such as soft tissues. Advancements in 3D reconstruction, notably through NeRF, have demonstrated great potential in solving these challenges. NeRF generates high-quality virtual representations of complex structures by training a neural network to describe the density and color information for each point of the scene in a continuous manner, even in challenging conditions [1, 2, 3]. For the clinical setting, a specialized scanning system is required to capture the precise datasets needed for NeRF-based and other RGB/RGB-D reconstruction methods. To meet this need, we present a novel surgical scanning system that reliably acquires NeRF-compatible data, enabling the rapid, high-quality digital twins of resected soft tissues for real-time surgical planning and post-operative analysis [4, 5].

Description of the New Method or System

We present a semi-automatic scanner platform designed for the rapid acquisition of RGB images, depth images, and point clouds, all with precisely known camera positions. The platform is further enhanced with NeRF-based algorithms to enable high-precision 3D reconstruction of soft tissues. The system integrates an RGB-D camera (Intel RealSense D405), a rotary stage (Standa 8MR190-2-28, 0.01° resolution), an LED lighting system, and a 3D-printed camera arm (Figure 1). The polylactic acid (PLA) camera arm allows the positioning of the pitch angle of the camera between 0° and 90° (5° resolution) and the positioning of the camera distance from the object between 60mm and 130mm (10mm resolution). The rotary stage automatically sets the yaw angle, whereas the pitch angle of the camera axis is adjusted manually on the camera arm. The soft tissues selected for macroscopic description vary in size and shape, ranging from 15mm to 65mm, and the adjustable camera distance ensures the preservation of their key features. The reconstruction process consists of (i) calibration phase with a ChArUco[1] marker-based camera pose detection, (ii) acquisition of RGB-D and point clouds phase, (iii) post-processing phase and (iv) training phase. The neural network was trained to create an implicit representation of the scene using Nerfstudio [4]. Nerfstudio also includes a tool for converting implicit representations to explicit formats, allowing for the export of point clouds. The training was performed on an NVIDIA T4 GPU within a Google Colab environment. The

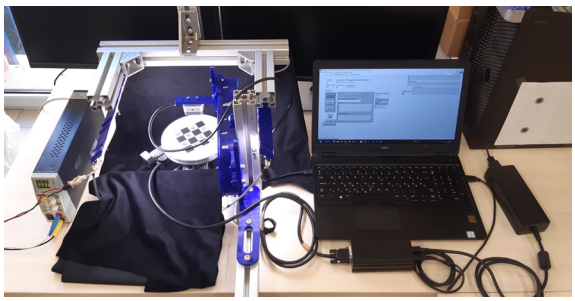

Fig. 1: Semi-Automatic RGB-D Scanning System.

[1] https://docs.opencv.org/4.x/d0/d3c/classcv_1_1aruco_1_1CharucoBoard.html

use of the Nerfacto method for the training of a single scene, comprising 200 images, typically required approximately 20 minutes, with a maximum of 10,000 iterations.

Results

To evaluate the effectiveness of both the scanning system and the algorithm, their outputs were benchmarked against a reference standard through two separate validation methods. The depth accuracy of the Intel RealSense camera was assessed at six sample pitch angles on the camera arm (25°, 35°, 45°, 60°, 75° and 90°), at multiple yaw angles for each pitch angle. This was repeated for three sample camera elevations (70mm, 90mm and 130mm). During these acquisitions, the ChArUco board was placed on the rotary stage with its width and height dimensions serving as the control parameter of the depth accuracy. The pattern real dimensions were measured with a caliper and compared with the same measurements carried out with CloudCompare[2] on the raw acquired point clouds. The average absolute distance error and the average standard deviation across all camera distances were 0.24mm and 0.354mm, respectively, which align with the manufacturer's submillimeter accuracy claim. The accuracy of the NeRF-based reconstruction method was assessed through the acquisition of 200 images, each at distinct yaw and pitch angles and camera elevation. The dataset comprised two objects with known geometries, specifically 30x30x30mm cubes made of aluminum and delrin. The NeRF reconstruction of each reference object was then compared to its original CAD model. As before, the mean distance error of each point in the point cloud was compared to the ground truth geometry of the CAD models with the CloudCompare software (see Figure 2). After a thorough tuning phase to optimize the scanning and reconstruction pipeline, the average mean distance error was 0.19 mm, and the standard deviation was 1.094 mm. The reconstruction pipeline proposed

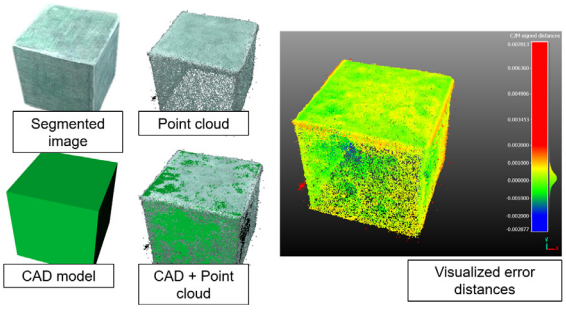

Fig. 2: Aluminum reference cube with mean error distances of the individual points.

[2]CloudCompare (version 2.13) [GPL software]. (2024). Retrieved from `http://www.cloudcompare.org/`

above was evaluated with resected animal soft tissues. An example is presented in Figure 3 to the reconstruction of an ex-vivo animal specimen. During the acquisition, the petri dish containing the specimen was placed on the rotary stage. Although no ground truth geometry was available for the soft tissue for thorough comparison, it is apparent that the reconstruction preserved the main geometric features while providing an accurate representation of the surface texture. Such tests revealed that the collected data can be effectively used for RGB and RGB-D based NeRF reconstruction methods and state-of-the-art 3D reconstruction methods. Therefore, the scanning system can be effectively utilized to collect data sets for the macroscopic description of ex-vivo soft tissues, with dimensions ranging between 15mm and 65mm [6]. Future work will focus on fine-tuning the presented algorithms to improve their performance and efficiency. This will include conducting comprehensive testing and parameter adjustments. Additionally, there will be an emphasis on gathering extensive and diverse datasets to enhance the accuracy, reliability, and generalizability of the system across various real-world scenarios.

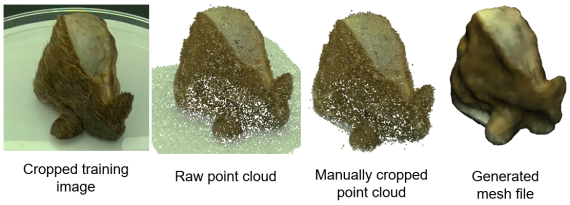

Fig. 3: Example of a 46x24x31mm canine mammary neoplasm, its reconstructed point clouds and generated mesh file.

References

[1] B. Mildenhall, P.P. Srinivasan, et al., Nerf: Representing scenes as neural radiance fields for view synthesis, *Commun. ACM*, 65, 99–106, 2021; doi: 10.1145/3503250

[2] A. Tewari, J. Thies, et al., Advances in neural rendering, *Comput. Graph. Forum*, 41, 703–735, 2022; doi: 10.1111/cgf.14507

[3] A. Yu, V. Ye, et al., pixelnerf: Neural radiance fields from one or few images, *CVPR*, 4578–4587, 2021; doi: 10.1109/CVPR46437.2021.00455

[4] M. Tancik, E. Weber, et al., Nerfstudio: A modular framework for neural radiance field development, *SIGGRAPH*, 1–12, 2023; doi: 10.1145/3588432.3591516

[5] A. Suvarna, R. Vempati, et al., DeltaAI: Semi-Autonomous Tissue Grossing Measurements and Recommendations using Neural Radiance Fields for Rapid, Complete Intraoperative Histological Assessment of Tumor Margins, *bioRxiv*; doi: 10.1101/2023.08.07.552349

[6] F. Auletta, M.C. Mezei, I. Benedetti, G. Rossi, A. Gavazza, C.M. Oddo, Digital Dataset of Formalin-Fixed Animal Tissues for Macroscopic Pathology: Image and Ultrasound Data, *in preparation*