2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

OCT-DeformNet: Optical Coherence Tomography Guided Biological Tissue Shape Prediction for Robot Palpation in Microsurgery

Guangshen Ma, PhD[*,†,1], Tianhao Qin[*,1], Jiawei Liu[1] and Mark Draelos, MD, PhD[1,2]

***Abstract*— In medical robotics, biological shape deformation resulting from arbitrary tool-tissue interaction commonly occurs and motivates the need in microsurgery to predict the shapes of tissue structures. However, handling deformation is challenging due to the lack of generalized prediction models for various surgical scenarios, complex tissue properties, and different surgical tool geometries. Limited intraoperative sensors to observe microlevel deformations further underscore this difficulty. This paper proposes a novel geometric data-driven framework that uses only the robot palpation tooltip movement and a pre-deformed surface to predict the tissue deformation by using the optical coherence tomography (OCT) sensor. A multilayer perceptron model is trained to learn tool-tissue physics and predict the shape from the given robot-tool configurations represented as orientations and displacements. We conducted realistic experiments to verify the models using phantoms of various stiffness and three *ex vivo* tissue types, with average prediction errors of approximately 0.15 mm and 0.52 mm respectively. This OCT-guided data-driven platform enables micro-scale palpation data collection and model training, and is broadly applicable to soft-tissue research in biomedical engineering and surgical robotics.**

I. Introduction

Tissue deformation can normally occur from many mechanisms intrinsically and extrinsically in biological systems. For example, in retinal surgery, an accurate shape prediction model can assist in intraoperative visualization and surgical planning to track tissue deformation resulting from the tip of the needle [1], [2]. The complex mechanics and microscopic scale of these nonlinear deformations present significant challenges in microsurgical environments, where surgeons find it increasingly difficult to anticipate tissue response prior to physical manipulation. Thus, accurately modeling and managing the behavior of such delicate tissues becomes an essential problem in robotic microsurgery.

Soft tissue surfaces can exhibit complex three-dimensional (3D) morphological behaviors governed by tool geometry, inherent mechanical properties, and local surface conditions [3], [4], [5]. For example, in teleoperated microsurgery, robotic palpation and interpretation of deformations are critical to compensate for surgeons' lack of direct haptic feedback [6]. While research has advanced for the topics of subsurface abnormality detection [7], tumor reconstruction [8], segmentation [9], and stiffness-based classification [10], achieving accurate shape prediction remains a challenge due to significant inter- and intra- tissue variability under dynamic tool-tissue interactions [11]. Thus, the primary challenges in robotic soft-tissue palpation can be categorized as follows: 1) modeling complex tool-tissue interaction properties; 2) characterizing tooltip dynamics under various orientations and displacements; 3) selecting intraoperative sensors capable of high-fidelity deformation capture; and 4) developing adaptive modeling frameworks suitable for diverse surgical scenarios.

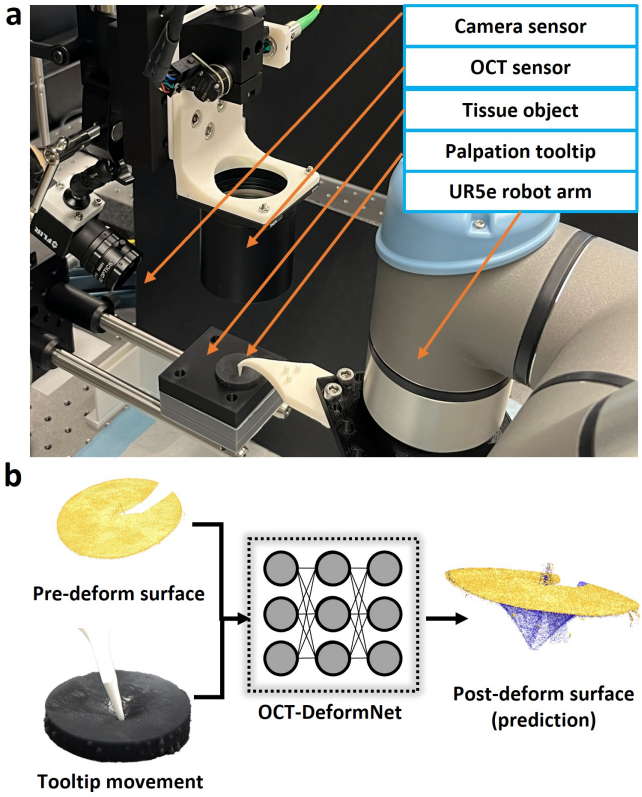

Fig. 1: OCT-guided tissue shape prediction from one-time robot palpation. (a) Data collection platform for OCT-guided robot palpation. (b) OCT-DeformNet to map a tooltip movement and a pre-deformed surface to a post-deformed surface.

* denotes equal contributions.
† Corresponding author: guangshe@umich.edu.
[1]GM, TQ (now at WPI), JL and MD are with the Department of Robotics, University of Michigan, Ann Arbor, MI USA.
[2]MD is also with the Department of Ophthalmology and Visual Sciences, University of Michigan Medical School, Ann Arbor, MI USA.
This work was supported by the National Institute of Biomedical Imaging and Bioengineering and the Office of the Director of the National Institutes of Health under award number DP2 EB039104 and by the University of Michigan. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

A. Challenges of robot palpation for tissue shape prediction

Complexity of robot tooltip movements in palpation: The complexity of tissue deformation in robotic palpation arises from the nonlinear, real-time, and the stochastic nature of tool-tissue interactions [12]. Variations in tooltip kinematics, including position, orientation, and tool geometry, can induce various types of deformation patterns. These complexities hinder the generalization of deformation models across diverse surgical instruments and robotic control strategies.

Limitations of contact sensing: To provide accurate tissue prediction models, sensory systems are important in predicting tissue deformation by providing multimodal data that inform the robot control system [10]. Most prior work focused on the use of force and tactile feedback to develop prediction models [13], [14], [5]. For example, tactile sensors observe fine-grained local deformations, while force feedback provides larger-scale resistance and kinesthetic sensations that allow a user or robot to understand and control the overall change in shape [15]. However, integrating force and tactile sensors directly into surgical tools is difficult. In minimally invasive surgery, sensors must be miniature, sterilizable, and robust for clinical use, which can often be neither practical nor cost-effective [16]. In addition, this sensory feedback remains insufficient for developing accurate, environment-specific models because indirect data cannot fully reconstruct complex 3D deformations [17]. Furthermore, indirect sensing often introduces prediction inaccuracies driven by complex tool-tissue interactions [12], hindering model generalization across heterogeneous anatomical structures.

Challenges with tissue deformation simulation: Predicting soft tissue deformation remains challenging due to inherent tissue complexities and unpredictable surgical dynamics, which frequently result in model inaccuracies and limited generalizability [18]. Although physics-based approaches, such as the Finite Element Method (FEM), offer high accuracy through mesh discretization, they often rely on idealized isotropic assumptions and suffer from high computational costs, making it difficult for real-time applications [19], [20]. The method of Position-Based Dynamics (PBD) provides an efficient alternative by modeling tissue as a system of interacting particles to simulate elasticity and tearing [21], [22], [23], [24]. However, both methods struggle to capture anisotropic or nonlinear behaviors and complex particle interactions, particularly as computational costs scale prohibitively with data volume and scene complexity. Nevertheless, our proposed method was inspired by the idea of discretization of points on the surface, which allows a data-driven approach to model the deformation behavior of each point directly from tissue sensor data.

Strengths of data-driven methods: Recent research in data-driven methods for modeling soft tissue deformation offers significant benefits, primarily by overcoming the limitations of traditional physics-based approaches and providing greater generalization capability for complex surgical scenarios [25], [18], [14]. Unlike conventional physics-based approaches, data-driven methods can be directly applied for complex tissue models without requiring explicit assumptions about anisotropic properties that are difficult to model accurately with standard formulas. In addition, recent advances in deep learning technologies significantly accelerate and improve the precision of soft tissue deformation modeling by offering faster real-time predictions and capabilities of handling complex real-world data [26], [27]. However, these end-to-end models require raw sensory data to train a large-scale model that can be computationally expensive [28].

B. OCT-guided data-driven modeling for shape prediction

Unique role of Optical Coherence Tomography: The limited capabilities of intraoperative sensors in microsurgery require a high-fidelity perception system to capture microstructure deformation. Conventional 3D sensors at the surface-level, such as the RGB-D depth camera, stereovision, 3D structured light, or laser scanning, are limited to capturing only the superficial layers of the surface [29]. These sensors also pose challenges when performing high-resolution scanning, mainly resulting from fundamental physical and technical constraints. This motivates the use of optical coherence tomography (OCT), a non-invasive imaging modality that uses low-coherence interferometry to detect reflected light and generate depth-resolved tissue maps with micrometer-level precision. By capturing volumetric information, OCT is well suited to measure micro-deformation induced by palpation [30], [31], and recent advances in OCT-guided surgery have stimulated algorithm development to handle dynamic surgical environments under tissue deformation [32].

OCT-based data-driven shape prediction model: In light of the challenges with modeling approaches and sensor limitations in capturing tissue microstructures, we develop "**OCT-DeformNet**", a novel data-driven framework to solve the shape prediction problem in robot palpation. By combining OCT-derived data, position-based dynamics (PBD)-inspired modeling strategies, and deep learning methods, we propose a novel geometric data-driven approach to map tooltip movements and initial surface shapes to the resulting post-deformed states. This framework discretizes the tissue into individual "particle-agents" and uses a multi-layer perceptron (MLP) to predict displacement vectors based on tooltip movements. By processing local geometric features and tooltip motion as inputs, the MLP estimates the displacement vector for each particle agent. The collective displacement of these particles reconstructs the post-palpation surface morphology, which can be used to effectively predict the surface deformation patterns directly from OCT data. In summary, the research contributions are:

- A robust OCT data collection framework for palpation-induced tissue deformation that captures a high-dimensional tooltip workspace across diverse orientations, displacements, and heterogeneous tissue types.
- A novel geometric data-driven model that predicts post-deformation surfaces using only tooltip movements and initial surface morphology. This multilayer perceptron-based (MLP) framework enables real-time inference for

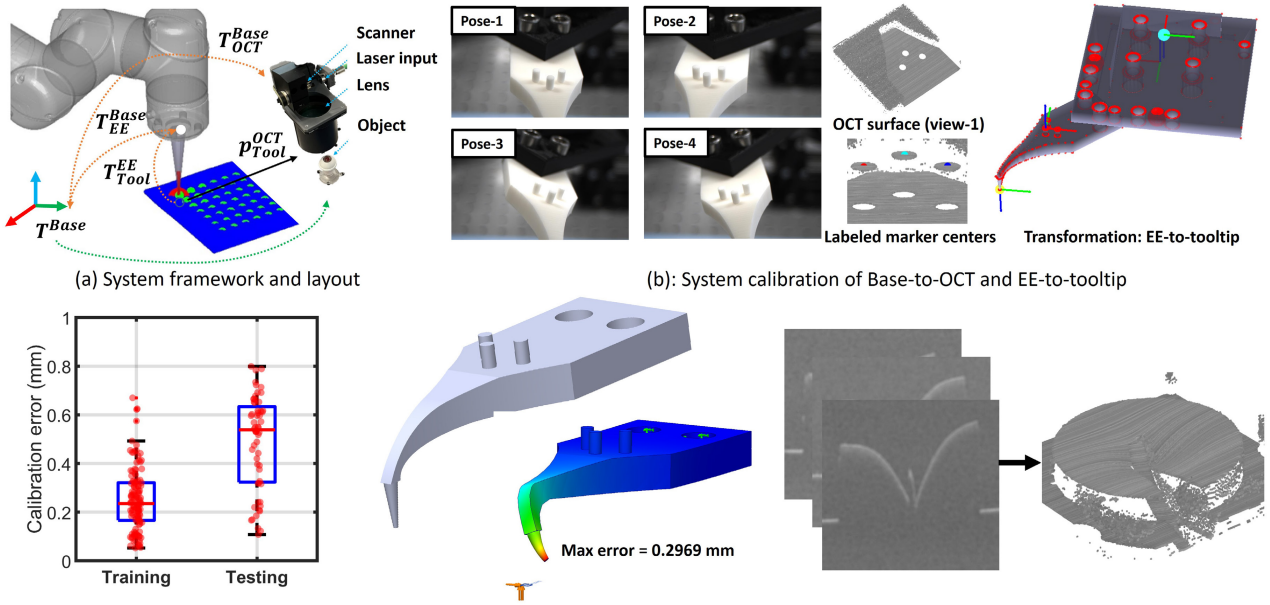

Fig. 2: (a) System configuration and the calibration model. (b) Overview of various marker poses for system calibration, the 3D reconstruction of the marker point cloud, and the transformation of the robot-EE to the tooltip. (c) Error boxplots of the calibration errors from the training (fitting) and the testing marker poses (different from the fitting model). (d) Finite element analysis for the maximum tool deformation with a 12N force applied at a 45-degree tilting angle (deformation enlarged 30 times than the actual deformation for visualization). (e) 3D surface reconstruction from sequential OCT B-scan images.

surgical planning and simulation in OCT-guided microsurgery without requiring prior knowledge of tissue properties or tool geometry.

- A validation approach on tissue phantoms and *ex vivo* tissues with live OCT data and real robot palpation.

II. Methods

A. System hardware

The system incorporates an UR5e robot arm (Universal Robots) with a surgical tool attached to the end-effector for robot palpation, a deformable tissue (phantom or *ex vivo* tissue), a tabletop OCT scanner and a color camera (Fig. 1).

1) OCT engine: We built a custom $1064\,\mathrm{nm}$ swept-source OCT engine with a Mach-Zehnder topology fiber interferometer and captured data using a balanced photoreceiver (Thorlabs) and a 1 GS^{-1} digitizer (ATS-9364; AlazarTech) using dual-edge sampling. For data collection, we used a $400\,\mathrm{kHz}$ source (Thorlabs) with an axial resolution of approximately 5 μm in air. The axial pixel pitch (for surface reconstruction) was calibrated as approximately 30 μm (after image resizing). The OCT covers a scanning region of approximately $37.4 \times 37.4\ mm^2$ area with a depth of about 10.2 mm, and a lateral resolution of about 0.11 mm. Each OCT volume takes approximately 1.0 seconds to acquire.

2) Robot tooltip design: A tool was attached to the robot end-effector to cover the palpation orientations ranging from $[-25, 0]$ degrees in tilt and $[-30, +30]$ degrees in sweep, both defined with respect to the robot base frame. The tool was designed with a curved geometry to minimize vision occlusion and ensure tool position accuracy even under large palpation forces. The tooltip was designed with 10mm length and 1.0 mm thickness. Three fiducial markers were attached in the middle of the tool to estimate the EE-to-tooltip transformation (Fig. 2b). The tool was 3D printed with glass-reinforced resin (Rigid 10K; Formlabs) to ensure exceptional stiffness and accurate manufacturing. A finite element analysis (FFEPlus iterative solve FEA solver) was performed to estimate the maximum deformation under a force of 12N with a 45-degree orientation towards the tip center (Fig. 2d). The maximum deformation was 0.30 mm.

3) OCT surface reconstructions: The OCT sensor collects sequential cross-sectional (B-scan) images of the entire scanning space. Each image was smoothed with a Gaussian filter for noise suppression, and the superficial layer was segmented by extracting the maximum-intensity pixel along each column. The stitching of the segmented B-scan image formed a surface (Fig. 2e).

B. System calibration

We performed a system calibration to estimate the Tool-to-Base and OCT-to-Base transformation (robot base frame as the world frame). We adjusted the conventional hand-eye calibration approach [33] and used three calibration markers to formulate a chessboard pattern attached to the tool (Fig. 2b). This formulates an optimization problem to calculate the optimal Tool-to-EE (T_{Tool}^{EE}) and OCT-to-Base

(T_{OCT}^{Base}) transformation (Fig. 2a):

$$\begin{aligned} F &= f(\ T_{OCT}^{Base};\ T_{Tool}^{EE};\ T_{EE}^{Base}(q_{calib})\ ) \\ &= T_{OCT}^{Base}\ p_{marker}^{OCT} -\ T_{EE}^{Base}(q_{calib}) T_{marker}^{EE}\ p^{marker} \end{aligned} \tag{1}$$

Where p^{marker} is the homogeneous vector and $T_{EE}^{Base}(q_{calib})$ is the Base-to-EE transformation with the joint configuration q_{calib}. The $q \in \mathbb{R}^6$ is a vector of joint variables. For the calibration, the robot-controlled tools were moved to various poses that covered translations and orientations associated with OCT scans. We obtain the cost function as:

$$\min_{T_{OCT}^{Base},\ T_{marker}^{EE}} \quad \sum_i || \ F_i(\ T_{OCT}^{Base};\ T_{marker}^{EE})\ ||_2^2 \tag{2}$$

The p_{marker}^{OCT} is the homogeneous position vector (center labeled from the OCT surface reconstruction) and p^{marker} is described by the marker coordinate (in the {marker} frame). We used the MATLAB optimization solver with the "Levenberg-Marquardt" algorithm to solve Eq. 2. The origin of the tooltip frame was determined by applying the optimal T_{marker}^{EE} and marker-to-tip transformation (obtained from the CAD model), and the orientations of the three axes are the same as the robot-EE (Fig. 2b). Thus, the information associated with tooltip movement and tissue deformation can be described in a single frame to train the prediction model.

C. MLP forward prediction with OCT-DeformNet

We implemented a fully-connected multilayer perceptron (MLP) framework to serve as the shape predictor. While Graph Neural Networks (GNNs) can capture tissue topology through message-passing layers, this comes at significant computational cost that hinders real-time performance [34], [35]. Instead, MLP provides a computationally efficient and lightweight alternative for tasks where inference speed is important in OCT-guided microsurgery. Thus, this study uses the MLP to demonstrate model feasibility. The MLP consists of four fully connected layers with 128, 128, 64, 32 neurons in each successive layer. The first half of the MLP input incorporates the tooltip movement, which is represented as the palpation vector from the palpation center to the target tooltip position (Fig. 3b). Notably, this palpation vector represents both the tooltip displacement and the orientation. The second half of the MLP input includes the tissue states by which the local geometry of the tissue relative to the center point can be described by a point-to-center displacement vector (Fig. 3b). Given a robot palpation center point and a pre-deformed surface point cloud, we have:

$$\Delta d_{output} = f_{MLP}(\ \Delta X_{robot},\ \Delta X_{tissue}). \tag{3}$$

Where $\Delta X_{robot} \in \mathbb{R}^3$ and $\Delta X_{tissue} \in \mathbb{R}^3$ represent the tooltip movement and the tissue state. $\Delta d_{output} \in \mathbb{R}^3$ denotes the displacement vector from the predeformed surface point. For the updated displacement, we have:

$$p_{k+1} = p_k + \Delta d_{output}, \tag{4}$$

where p_k and p_{k+1} denote the pre-deformed and post-deformed surface point updated by the displacement vector predicted from the MLP. The summation of all the points formulates an updated post-deformed surface denoted as a point cloud object.

1) Loss function: The shape prediction is modeled as a regression problem where the predictor is represented by the MLP. We applied the Mean Squared Error (MSE) loss function to calculate the average of the squared differences between the predicted and the true displacement vectors,

$$\mathcal{L}_{MSE} = \frac{1}{N} \sum_{i=1}^{N} ||\mathbf{y}_i - \hat{\mathbf{y}}_i||_2^2, \tag{5}$$

where $\mathbf{y}_i \in \mathbb{R}^3$ denotes the predicted displacement vector and $\hat{\mathbf{y}}_i \in \mathbb{R}^3$ represents the labeled vectors extracted from the OCT data through preprocessing in Fig. 3a. This model maps the movement of the robot and a pre-deformed surface directly to the post-deformed surface without providing the knowledge of tissue properties and tool configurations.

MLP training: The MLP model was trained with Nvidia RTX A2000 12GB GPU on Windows 10 with the Pytorch deep learning framework. We trained the model with 60 epochs and selected the last one for the testing experiments.

D. OCT data collection and analysis

To train OCT-DeformNet, we developed a data collection platform to achieve palpation in arbitrary patterns. A predeformed surface was first scanned by the OCT sensor. The center of the surface was selected as the palpation center.

1) Local surface geometry: We define a local coordinate frame at the palpation center following the same axis orientations as the robot base frame (Fig. 3a). This standardized definition enables the model to be used in different robot coordinate systems. To mitigate data bias arising from varying deformation magnitudes across the tissue surface, we defined a fixed region of interest (ROI) around the palpation center that captures the dominant deformation field.

2) Tooltip movement space (robot inputs): We define a sequence of tooltip movements to achieve robot palpation covering tilting angles from -25 to 0 degrees and sweeping angles of -30 to 30 degrees (around the vertical axis in the local frame) at the palpation center (Fig. 3a). For each orientation set, the tool was controlled to perform a penetration displacement ranged from 0.0 to 8.0 mm (limited by the maximum imaging depth of the OCT sensor).

3) Post-deformed surface processing: With a stop-and-scan strategy, the robot induced palpation-based tissue deformation and remained stationary during OCT acquisition. As the tool caused vision occlusion within the OCT field-of-view and produced a well-described complex-conjugate artifact, we transformed the points of the tool (from CAD models) to the base frame using robot kinematics and calibration information, as shown in Fig. 3. The point cloud within the convex hull formulated by the tool boundary was removed to eliminate vision occlusion [36].

4) Local tissue geometry (tissue inputs): As the initial tissue shape significantly influences how it deforms in response to pressure, local geometric information is needed to serve

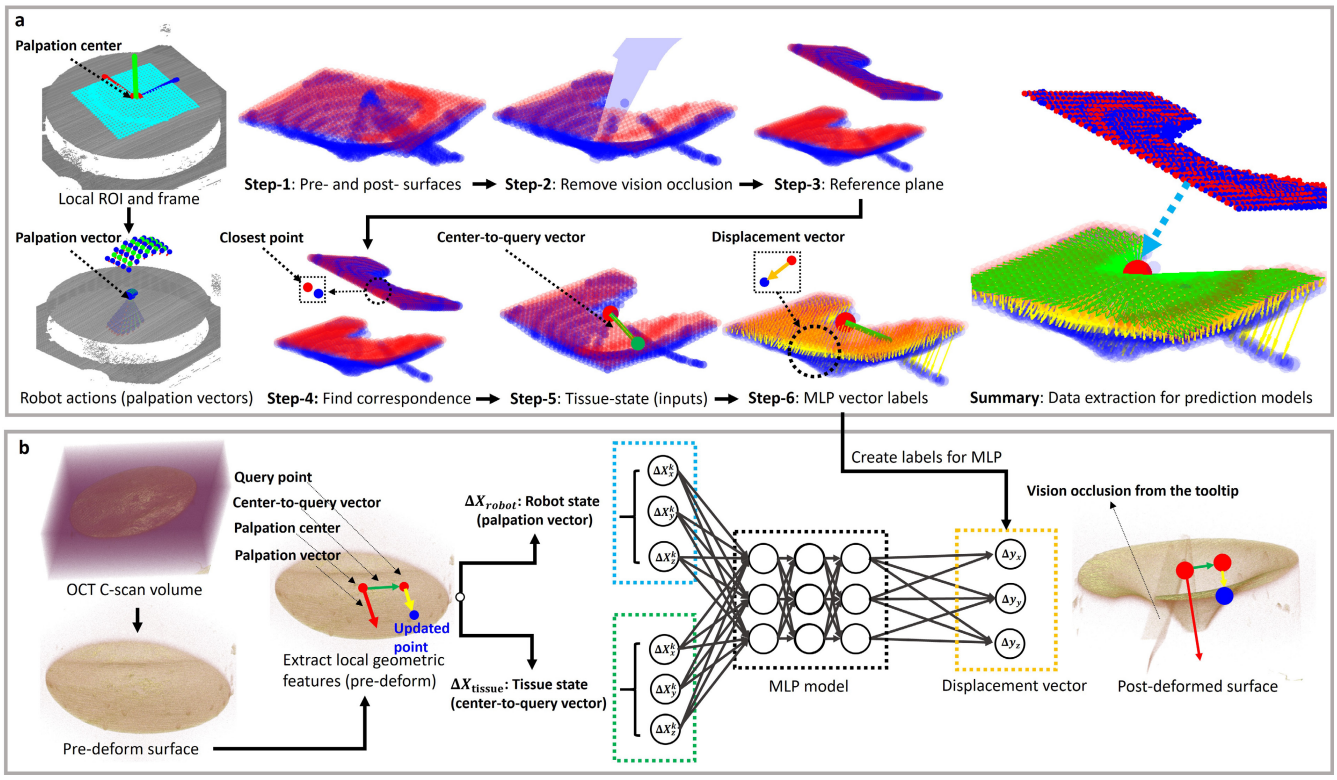

Fig. 3: Summary of the data analysis workflow and the MLP model. (a) The analysis workflow for calculating the MLP inputs and labels. The local frame is initially defined around the palpation center and the tooltip movement is described by the palpation vector encoding both the displacement and the orientation. Step-1 to Step-6 shows the workflow to remove the pointcloud from the vision occlusion (some pointcloud outliers would remain due to the offset of the system calibration), calculation of the tissue states, and the representations of displacement vectors (labels). (b) The MLP model maps the inputs of a pre-deformed surface and a tooltip movement described by the palpation vector to the post-deformed surface.

as inputs to the MLP. Although existing solutions, such as variants of the graph neural network and 3D convolutional neural networks, have been proposed to solve this problem [34], encoding spatial information in the neural network remains a challenging problem [37]. Additionally, calculating the relation between a query point and other data points can be computationally expensive due to the complexity of the distance metric, the size of the surface data, and the lack of efficient indexing [38]. Thus, we used the point-to-center vector to represent the geometric information between a query point and the palpation center (Fig. 3a step-5).

From a practical standpoint, our framework treats each query point within the point cloud as an independent agent. While this approach currently does not capture complex global tissue geometries, we adopt the reasonable assumption that local surgical sites often approximate a planar surface within a small region-of-interest (versus a large-area surface).

5) MLP model outputs and labels: For each data point from the surface, a single point in the pre-deformed surface can be matched to multiple points in the post-deformed surface. Thus, we developed a reference plane that follows the same palpation orientation to calculate the displacement vector. The points of the pre- and post- deformed surfaces were first projected to the reference plane (Fig. 3a step-4). The point correspondences were built using the k-Nearest Neighbors (KNN) algorithm in MATLAB, which was used to calculate the displacement vector (from pre- and post- point correspondence). These vectors served as MLP outputs and training labels in the datasets (Fig. 3a step-6).

III. Experiments

A. System calibration verification

For system calibration, we collected OCT scans of the fiducial markers with 42 poses to cover the orientation and positions. For training (fitting), Eq. 2 was solved to estimate the optimal Tool-to-EE and OCT-to-Base transformation. The reprojection errors were reported as 0.25 ± 0.13 mm. For testing, we collected a new dataset with 20 robot poses and OCT scans. The optimal Tool-to-EE and OCT-to-Base transformation was applied to calculate testing errors, with an average value reported as 0.49 ± 0.19 mm (Fig. 2c).

B. MLP model training and inference performance

The MLP training performances are shown in Fig. 4c with three different random seeds. We selected the default random seed to report the model performance result in Fig. 4d. In addition, we evaluated the MLP models across

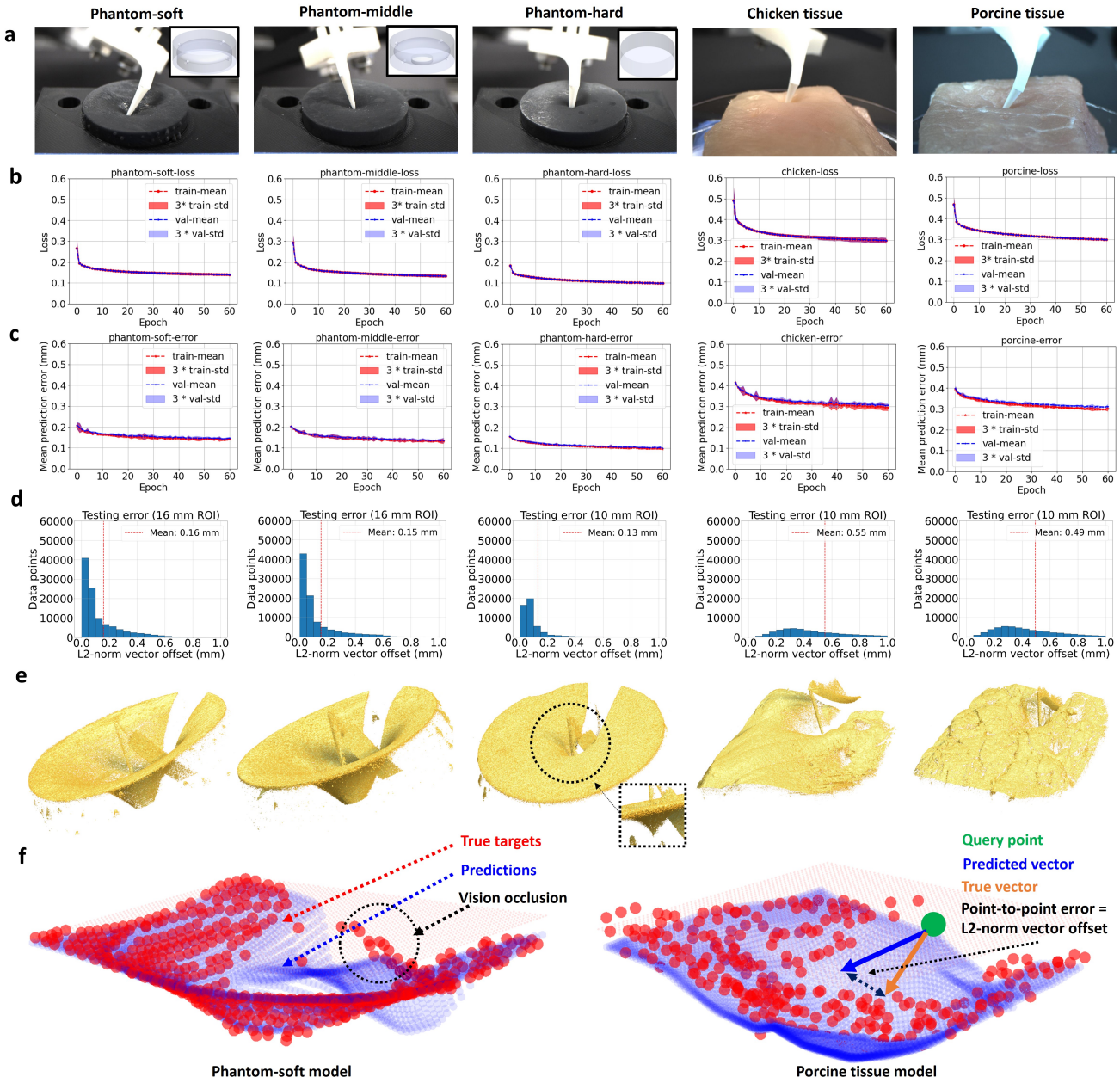

Fig. 4: Summary of phantom and *ex vivo* studies. (a) The tissue models from left to right are 1) soft stiffness phantom, 2) middle stiffness phantom, 3) hard stiffness phantom, 4) chicken tissue, and 5) porcine tissue. (b) The loss curves of the MLP training (with three random seeds). (c) The error curves of the MLP training (with three random seeds). (d) Histogram of the L2-norm vector offsets (equivalent to the updated point offsets between the prediction and the true ones). The default random seed is selected. (e) Represented surface measurements of each tissue type under 8 mm palpation. (f) Deformation prediction of phantom-soft and porcine models (red: true targets, blue: prediction) and the visualization of the vector offset and point-to-point error (pre- and post- difference). A portion of the point cloud was removed from vision occlusion and thus the MLP model cannot achieve accurate prediction at the region obscured by the tool.

various point cloud sizes (densities) for inference speeds, using five tissue models (pre-deformed surface) reconstructed from post-processed OCT data (Table. I). The pre-processing incorporates the procedure of organizing the pointcloud data to the standardized format of MLP inputs.

C. *Error evaluation metrics*

For each tissue object, we collected a testing dataset from various palpation configurations (for both the phantom study and *ex vivo* model) and various surfaces (*ex vivo* only). Palpation vectors, center-to-query vectors (tissue geometry), and displacement vectors were calculated in the testing

dataset. The trained MLPs were applied to calculate the predicted vectors. As an MLP regression problem, the error was computed as the L2-norm between the predicted and the true vectors from testing dataset. This L2-norm error is also equivalent to the distance between the predicted and true points after deformation. The same metric applied for both the phantom and *ex vivo* studies.

TABLE I: MLP model inference time (average with 1000 repeated measurements) with various pointcloud densities (testing GPU: NVIDIA RTX A2000 12GB).

	Points (N)		**Prediction (ms)**		**Preprocessing (ms)**	
Density	Mean	SD	Mean	SD	Mean	SD
Low	9877	839	0.70	0.47	0.22	0.42
Medium	39005	3780	1.77	0.55	1.01	0.40
High	109087	4586	4.95	1.39	3.52	0.53

D. Experiment-1: Phantom study

1) Phantom designs: We performed the phantom study to evaluate OCT-DeformNet for shape predictions. We developed three phantom models manufactured with Silicon 40A (Formlabs) material. We designed an empty media layer to produce various stiffness levels, described as phantom-soft (large empty media with 2.0 mm thickness), phantom-middle (empty media with 6.0 mm thickness), and phantom-hard (no empty media). The CAD design is shown in Fig. 2a.

2) Training and validation: We formulated the training datasets through robot palpation within a tooltip movement space. We applied orientations with tilting angles ranged from -24 to 0 with a step size of 6 degrees, and sweeping angles ranged from -30 to +30 with a step size of 12 degrees. Once the orientation vector was determined by the tilting and sweeping angles, the palpation distance was selected with a range from 0.0 to 8.0 mm and a step size of 0.8 mm. The scanning regions were selected as rectangular region of interest (ROI) of 16×16 mm for the soft and middle phantom and ROI of 10×10 mm for the hard phantom (due to very small deformation). The pointcloud was first downsampled (grid step: 0.5mm) to minimize the number of points within a small ROI (avoid overfitting of the MLP) and an approximately 600 data points were formulated for each scan. The MLP training results are shown in Fig. 4b-c.

3) Testing studies: With the same predeformed surface, we collected a separate testing dataset using the palpation settings of the tilting angles ranged from -21 to -3 with a step size of 6 degrees, the sweep angles ranged from -24 to +24 with a step size of 12 degrees, and the penetration distance ranged from 0 to 8.0 mm with a step size of 0.8 mm. The error statistics are reported in Fig. 4d. The average testing errors for the three phantom are reported as 0.16 mm with a standard deviation (SD) of 0.34 mm (soft), 0.15 mm with a SD of 0.36 mm (middle) and 0.13 mm with a SD of 0.32 mm (hard). The average point error contributes less than 10% of the maximum penetration depth.

E. Experiment-2: Ex vivo tissue study

1) Training and validation: We developed a similar OCT data collection method for various types of *ex vivo* tissue. Two types of *ex vivo* tissues (porcine, chicken) were used. The same workflow as in the phantom study was applied to collect data for the *ex vivo* tissue model, including the same configurations for tilting, sweeping, and penetration distance. Unlike the phantoms, palpating *ex vivo* tissue was more challenging since the surface did not always recover to its original, and therefore a single contact spot cannot be used repeatedly. Instead of a single palpation center, we resampled the surface with a fixed step size of 10 mm for multiple point candidates. Each point was only used for one tilting angle associated with all the sweeping ones. The scanning region for the *ex vivo* tissue model was a 10×10 mm square ROI.

2) Testing studies: We used a completely different tissue surface (unused during training) to formulate the testing dataset. The same setting was applied (as in phantom) for the testing of *ex vivo* tissue. The results show average errors of 0.55 ± 0.50 mm and 0.49 ± 0.42 mm for the chicken and porcine tissue models, respectively.

IV. Discussions and Conclusions

We proposed an OCT-guided data-driven method to predict soft tissue deformation under palpation. The experimental results in phantom and *ex vivo* tissue studies demonstrate the feasibility of using OCT-DeformNet for soft tissue deformation. The reduced accuracy from phantom to *ex vivo* tissue models is due to the inherent inhomogeneity of biological tissue and the variability of tissue response from the same palpation patterns. Additional complexities from the vision occlusion problem also contribute to errors due to partial data removal from the collected ones. Nonetheless, the prediction model with this sub-millimeter accuracy can be used in providing real-time visualization and surgical planning.

For potential clinical applications, the proposed OCT-guided tissue deformation prediction model provides a robust, data-driven alternative to traditional physics-based approaches in robot-assisted microsurgery, effectively addressing intra-class tissue variability where standard models often fail to bridge the pre-to-intraoperative gap. This framework facilitates the efficient development of patient-specific models through a streamlined "collect-train-deploy" workflow that captures realistic tissue deformation physics directly from intraoperative interactions, eliminating the need for prior knowledge of complex tissue mechanics.

In summary, we used an entirely data-driven method to model tissue deformation resulting from robot palpation. A multilayer perceptron network was implemented for shape prediction with inputs of tooltip movement and tissue geometry, and can predict the displacement vector from a given point on the surface. Future work will focus on optimizing the scanning strategy to improve the efficiency of OCT data collection, and the exploration of MLP model parameters to improve the learning performances.

V. Acknowledgments

The authors would like to thank Yilin Ma from the Hybrid Dynamic Robotics Lab for the suggestions of 3D printing.

References

[1] S. Pannek, S. Dehghani, M. Sommersperger, P. Zhang, P. Gehlbach, M. A. Nasseri, I. Iordachita, and N. Navab, "Exploring the needle tip interaction force with retinal tissue deformation in vitreoretinal surgery," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 16 999–17 005.

[2] P. K. Gupta, P. S. Jensen, and E. de Juan Jr, "Surgical forces and tactile perception during retinal microsurgery," in *International conference on medical image computing and computer-assisted intervention*. Springer, 1999, pp. 1218–1225.

[3] J. Wang, Z. Zhou, Y. Cui, J. Cui, L. Qin, T. Xu, Z. Yi, and X. Wu, "Dual autoencoder-based joint learning: Enhancing depth classification of hard inclusions in soft tissue for robotic palpation," *IEEE/ASME Transactions on Mechatronics*, 2025.

[4] S. Misra, K. Ramesh, and A. M. Okamura, "Modeling of tool-tissue interactions for computer-based surgical simulation: A literature review," *Presence*, vol. 17, no. 5, pp. 463–491, 2008.

[5] B. Ahn and J. Kim, "Measurement and characterization of soft tissue behavior with surface deformation and force response under large deformations," *Medical image analysis*, vol. 14, no. 2, pp. 138–148, 2010.

[6] M. Mir, J. Chen, A. Patel, M. R. Pinezich, B. A. Guenthart, G. Vunjak-Novakovic, and J. Kim, "A minimally invasive robotic tissue palpation device," *IEEE Transactions on Biomedical Engineering*, vol. 71, no. 6, pp. 1958–1968, 2024.

[7] A. Bhattacharjee, M. T. Loghmani, and S. Anwar, "Modeling and simulation of robotic palpation to detect subsurface soft tissue anomaly for presurgical assessment," *Journal of Engineering and Science in Medical Diagnostics and Therapy*, vol. 7, no. 3, p. 031002, 2024.

[8] R. Uppuluri, A. Bhattacharjee, S. Anwar, and Y. She, "Seebelow: Subdermal 3d reconstruction of tumors with surgical robotic palpation and tactile exploration," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 6961–6968.

[9] Y.-j. Lee, S.-w. Bang, J.-b. Hong, and S. Park, "Image-free tumor segmentation of soft tissue using a minimally invasive robotic palpation system," *IEEE Transactions on Biomedical Engineering*, 2025.

[10] K. A. Nichols and A. M. Okamura, "Methods to segment hard inclusions in soft tissue during autonomous robotic palpation," *IEEE Transactions on Robotics*, vol. 31, no. 2, pp. 344–354, 2015.

[11] J. Zhang, Y. Zhong, and C. Gu, "Deformable models for surgical simulation: a survey," *IEEE reviews in biomedical engineering*, vol. 11, pp. 143–164, 2017.

[12] M. Neidhardt, R. Mieling, M. Bengs, and A. Schlaefer, "Optical force estimation for interactions between tool and soft tissues," *Scientific reports*, vol. 13, no. 1, p. 506, 2023.

[13] D. Pinzon, S. Byrns, and B. Zheng, "Prevailing trends in haptic feedback simulation for minimally invasive surgery," *Surgical innovation*, vol. 23, no. 4, pp. 415–421, 2016.

[14] Y. Zhu, L. Abdulmajeid, and K. Hauser, "A data-driven approach for fast simulation of robot locomotion on granular media," in *2019 international conference on robotics and automation (ICRA)*. IEEE, 2019, pp. 7653–7659.

[15] A. M. Okamura, "Haptic feedback in robot-assisted minimally invasive surgery," *Current opinion in urology*, vol. 19, no. 1, pp. 102–107, 2009.

[16] J. Konstantinova, A. Jiang, K. Althoefer, P. Dasgupta, and T. Nanayakkara, "Implementation of tactile sensing for palpation in robot-assisted minimally invasive surgery: A review," *IEEE Sensors Journal*, vol. 14, no. 8, pp. 2490–2501, 2014.

[17] T.-N. Nguyen, M.-C. Ho Ba Tho, and T.-T. Dao, "A systematic review of real-time medical simulations with soft-tissue deformation: Computational approaches, interaction devices, system architectures, and clinical validations," *Applied Bionics and Biomechanics*, vol. 2020, no. 1, p. 5039329, 2020.

[18] H. Liu, Y. Han, D. Emerson, Y. Rabin, and L. B. Kara, "A data-driven approach for real-time soft tissue deformation prediction using nonlinear presurgical simulations," *PloS one*, vol. 20, no. 4, p. e0319196, 2025.

[19] X. Liang, F. Liu, Y. Zhang, Y. Li, S. Lin, and M. Yip, "Real-to-sim deformable object manipulation: Optimizing physics models with residual mappings for robotic surgery," in *2024 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2024, pp. 15 471–15 477.

[20] F. Liu, Z. Li, Y. Han, J. Lu, F. Richter, and M. C. Yip, "Real-to-sim registration of deformable soft tissue with position-based dynamics for surgical robot autonomy," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 12 328–12 334.

[21] M. Camara, E. Mayer, A. Darzi, and P. Pratt, "Soft tissue deformation for surgical simulation: a position-based dynamics approach," *International journal of computer assisted radiology and surgery*, vol. 11, no. 6, pp. 919–928, 2016.

[22] J. Pan, J. Bai, X. Zhao, A. Hao, and H. Qin, "Real-time haptic manipulation and cutting of hybrid soft tissue models by extended position-based dynamics," *Computer Animation and Virtual Worlds*, vol. 26, no. 3-4, pp. 321–335, 2015.

[23] Y. Guo, D. Kou, X. Zhang, L. Zhang, and K. Cao, "A real-time virtual liver simulator using position based dynamics with compliant constraints for medical applications," *Scientific Reports*, vol. 15, no. 1, p. 26557, 2025.

[24] J. Fang, L. You, E. Chaudhry, and J. Zhang, "State-of-the-art improvements and applications of position based dynamics," *Computer Animation and Virtual Worlds*, vol. 34, no. 5, p. e2143, 2023.

[25] Z. Zhai, E. Wang, and X. Chen, "Fast geometric deep learning for intraoperative soft tissue deformation estimation: Towards real-time ar guidance in liver surgery," *Medical Engineering & Physics*, p. 104409, 2025.

[26] M. Kim, G. Pons-Moll, S. Pujades, S. Bang, J. Kim, M. J. Black, and S.-H. Lee, "Data-driven physics for human soft tissue animation," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 4, pp. 1–12, 2017.

[27] R. D. Huff, F. Houghton, C. C. Earl, E. Ghajar-Rahimi, I. Dogra, D. Yu, C. Harris-Adamson, C. J. Goergen, and G. D. O'Connell, "Deep learning enables accurate soft tissue tendon deformation estimation in vivo via ultrasound imaging," *Scientific Reports*, vol. 14, no. 1, p. 18401, 2024.

[28] N. Gessert, M. Bengs, M. Schlüter, and A. Schlaefer, "Deep learning with 4d spatio-temporal data representations for oct-based force estimation," *Medical image analysis*, vol. 64, p. 101730, 2020.

[29] R. H. Taylor, A. Menciassi, G. Fichtinger, P. Fiorini, and P. Dario, "Medical robotics and computer-integrated surgery," in *Springer handbook of robotics*. Springer, 2016, pp. 1657–1684.

[30] D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, *et al.*, "Optical coherence tomography," *science*, vol. 254, no. 5035, pp. 1178–1181, 1991.

[31] W. Drexler and J. G. Fujimoto, *Optical coherence tomography: technology and applications*. Springer Science & Business Media, 2008.

[32] G. Ma, M. McCloud, Y. Tian, A. Narawane, H. Shi, R. Trout, R. P. McNabb, A. N. Kuo, and M. Draelos, "Robotics and optical coherence tomography: current works and future perspectives," *Biomedical Optics Express*, vol. 16, no. 2, pp. 578–602, 2025.

[33] R. Y. Tsai, R. K. Lenz, *et al.*, "A new technique for fully autonomous and efficient 3 d robotics hand/eye calibration," *IEEE Transactions on robotics and automation*, vol. 5, no. 3, pp. 345–358, 1989.

[34] D. Li, C. Lu, Z. Chen, J. Guan, J. Zhao, and J. Du, "Graph neural networks in point clouds: A survey," *Remote Sensing*, vol. 16, no. 14, p. 2518, 2024.

[35] R. P. McKennaa and J. Oyekan, "A perspective on open challenges in deformable object manipulation," *arXiv preprint arXiv:2602.22998*, 2026.

[36] F. Y. Yap, J. T. Bui, M. G. Knuttinen, N. M. Walzer, S. J. Cotler, C. A. Owens, J. L. Berkes, and R. C. Gaba, "Quantitative morphometric analysis of hepatocellular carcinoma: development of a programmed algorithm and preliminary application," *Diagnostic and interventional radiology*, vol. 19, no. 2, p. 97, 2013.

[37] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 652–660.

[38] W. Shi and R. Rajkumar, "Point-gnn: Graph neural network for 3d object detection in a point cloud," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 1711–1719.