

Feedback Matters: Augmenting Autonomous Dissection with Visual and Topological Feedback

Chung-Pang Wang^{*1}, Changwei Chen^{*1}, Xiao Liang¹, Soofiyan Atar¹, Florian Richter¹, Michael Yip¹

Abstract—Autonomous surgical systems must adapt to highly dynamic environments where tissue properties and visual cues evolve rapidly. Central to such adaptability is feedback: the ability to sense, interpret, and respond to changes during execution. While feedback mechanisms have been explored in surgical robotics, ranging from tool and tissue tracking to error detection, existing methods remain limited in handling the topological and perceptual challenges of tissue dissection. In this work, we propose a feedback-enabled framework for autonomous tissue dissection that explicitly reasons about topological changes from endoscopic images after each dissection action. This structured feedback guides subsequent actions, enabling the system to localize dissection progress and adapt policies online. To improve the reliability of such feedback, we introduce visibility metrics that quantify tissue exposure and formulate optimal controller designs that actively manipulate tissue to maximize visibility. Finally, we integrate these feedback mechanisms with both planning-based and learning-based dissection methods, and demonstrate experimentally that they significantly enhance autonomy, reduce errors, and improve robustness in complex surgical scenarios.

I. INTRODUCTION

Learning and adaptation are impossible without observational feedback. In human cognition, feedback is the core mechanism for mastering skills: from infants adjusting their first steps after a stumble, to expert surgeons refining their skill based on subtle visual cues. Robotic autonomy similarly depends on feedback, where sensing the environment and responding to unexpected changes are essential for robust and adaptive behavior. Surgical robotics is no different especially for sub-tasks that demands high precision and reactive capability to complex nonlinear changes, such as tissue dissection and suturing. Yet current surgical automation works lacks the reliable feedback-enabled autonomy required to detect and react to failures in real tissue environments.

Feedback mechanisms have been explored in surgical robotics to enable more adaptive and reliable autonomy. A prominent line of work has focused on tool tracking [1], [2] with the da Vinci Research Kit (dVRK), which has enabled applications such as visual servoing of surgical instruments [3]. Similarly, needle tracking [4] has been investigated to facilitate suturing subtasks, including the delicate hand-off of the needle between instruments. Tissue modeling and tracking [5], [6] from real-world sensor data provide dynamic estimates of tissue deformation and support safe manipulation [7] and retraction [8]. Other feedback

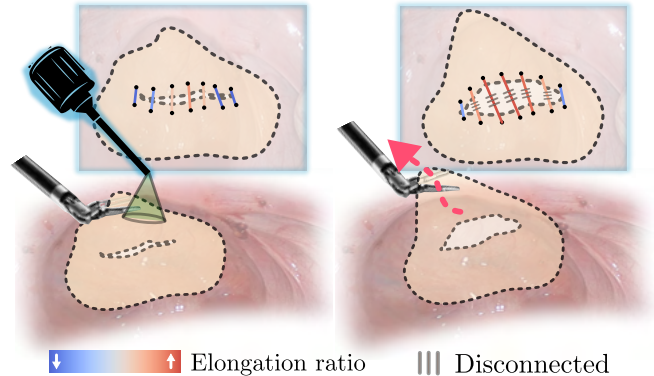


Fig. 1: Illustration of the importance of exposure maximization for tissue dissection outcome verification. When tissue is dissected, an estimation method is uncertain about tissue connectivity condition due to lack of stretching. A proper controller should not only stretch the dissection site, but also consider its visibility under the camera observation model.

mechanisms address error detection, such as methods for blood flow detection that leverage physics models and visual cues to identify bleeding events and guide blood suction [9]. More recently, learning-based methods [10], [11] have emerged for their improved precision and generalization ability in the presence of occlusion, noise, and modeling error. Learned policy internalizing sensor feedback through data-driven optimization. However, they lack explainability and theoretical guarantees, and their performance often degrades and produce unpredictable failures in the absence of large scale datasets, limiting its reliability in surgical settings.

Despite progress in feedback mechanisms for surgical automation, relatively little attention has been given to the high-precision surgical task of tissue dissection. Previously, several works have attempted to automate dissection through supervised controllers for precise 3D incisions on porcine tissues [12], [13], frameworks for boundary detection and tool tracking [14], and customized systems for tumor resection [15], [16]. Reinforcement learning has also been applied to automate electrocautery dissection from point cloud input [17], while other efforts target sensing and reconstruction of dissection sites [18], [19]. In parallel, simulation-based research has advanced modeling of tissue dissection [20], [21], providing important platforms for testing and training autonomous surgical systems. Yet, despite their contributions, none of these methods have explored incorporating robust feedback mechanisms. They automate execution given predefined goals but cannot verify task completion or adapt when dissection progresses differently than expected, which

^{*}These authors contributed equally to this paper.

¹ Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093, USA {chw120, chc165, x5liang, satar, frichter, yip}@ucsd.edu

is a critical shortcoming in real surgical environments. In addition, many critical feedback signals are not directly observable and must be actively maximize exposure through manipulation—for example, stretching tissue to verify that a cut has been fully completed. This tight coupling between perception and action underscores the need for feedback mechanisms that go beyond passively interpreting sensor data. Instead, they must also drive active interventions that expose hidden state information essential for safe, accurate, and autonomous dissection.

To address these challenges, we propose a feedback-enabled framework for autonomous tissue dissection. Our method explicitly reasons about topological changes of tissue from endoscopic image data after each dissection action, providing structured feedback to subsequent steps of an autonomous dissection agent. Furthermore, we introduce two visibility metrics of dissected areas, which quantify the reliability of the feedback based on tissue surface exposure to the camera and can be integrated into a dissection agent to improve the quality of the feedback. Finally, through extensive experiments, we demonstrate the versatility of our visual and topological feedback for tissue by integrating it with both planning-based and learning-based dissection methods, analyzing their limitations, and showing experimentally how feedback improves autonomy in each case.

II. PRELIMINARIES

Deformable Tissue Simulation: Extended Position-based Dynamics (XPBD) simulator has been used as dynamic models for tissue manipulation in previous works [3]. This work adopts it as the underlying dynamic model for exposure maximization control. XPBD represents a system with particles \mathbf{x}_t , robot control u_t and geometric constraints, iteratively solving particle states for constraint satisfaction. The dynamic equation can be expressed as

$$\mathbf{x}_{t+1} = f_{\text{xpbd}}(\mathbf{x}_t, u_t) \quad (1)$$

A piece of tissue is represented a triangular mesh $\mathbf{M} : \{\mathbf{v}, \mathbf{F}\}$, where \mathbf{v} is a set of mesh vertices and $\mathbf{F} : \{F_1, \dots, F_b\}$ denotes triangle faces. The tissue is simulated with vertices as particles, and with distance constraints defined on its edges, bending constraints defined on neighboring faces. The robot control $\mathbf{u}_t \in \mathbb{R}^3$ can achieve grasping and tissue retraction effect by simulating it as a boundary condition on position offsets of certain tissue particles.

Tissue tracking: Tissue tracking is an important technique for localizing and tracking anatomy over time. Tissue tracking methods are denoted as

$$p_j = \mathcal{Q}_i^j(p_i) \quad (2)$$

where i, j indicates time stamps or frames. The tracking method takes in coordinate query p_i at frame i and produces the corresponding coordinate p_j at frame j , describing the movement of p . In this work, we consider two of the most widely used approaches for tissue tracking: image keypoint tracking (IKT) ${}^{ikt}\mathcal{Q}$ and model-based tracking (MBT) ${}^{mbt}\mathcal{Q}$.

While both methods track point in the image frame (i.e. $\mathcal{Q} : \mathbb{R}^2 \mapsto \mathbb{R}^2$), their tracking mechanisms are different. IKT tracks points purely with flow with image feature, such as [22]. MBT tracks point based on a physics model's prediction, in this case, XPBD model. The full tracking procedures are defined as follow:

$$\begin{aligned} {}^{mbt}\mathcal{Q}_i^j(p_i, \mathbf{x}_{i:j}) : \hat{p}_i &= \text{raycast}(p_i, \mathbf{x}_i, \mathbf{M}, C) \\ k, \alpha, \beta, \gamma &= \text{barycentric}(\hat{p}, \mathbf{x}_i, \mathbf{M}) \\ \hat{p}_j &= \text{track}(\mathbf{x}_j, k, \alpha, \beta, \gamma) \\ p_j &= \text{project}(\hat{p}_j, C). \end{aligned} \quad (3)$$

In above equations, C represents a known model-view matrix. These procedure includes (1) performing ray-casting of an image point to back-project it onto the mesh surface, (2) using its barycentric coordinates to track the point after mesh deform, (3) project the 3D point back to the image frame. For simplicity, dependency of ${}^{ikt}\mathcal{Q}$ on image inputs, and ${}^{mbt}\mathcal{Q}$ on the mesh and camera model are omitted.

III. METHODS

The proposed method considers a task where a dissection goal is provided as an input by expert surgeons, and it aims at removing tissue connectivity around the given target. A dissection goal is parameterized as a polyline $\mathcal{G}^{(t)} = \{p_t^k \in \mathbb{Z}^2 \mid k = 1, \dots, M\}$ that lies on the pixel space on an endoscopic image. This representation is chosen as endoscopic displays are the most common interaction modality. We assume the dissection goal to be a low-curvature path and we denote $\bar{\mathbf{n}}_t, \bar{\mathbf{m}}_t$ to be the orthogonal and parallel unit vectors of the dissection goal $\mathcal{G}^{(t)}$. An intuitive understanding of tissue connectivity removal is that tissue connections do not intersect with the dissection polyline, \mathcal{G} . Its exact definition and metrics is detailed in Sec III-B.

This work particular concerns the completeness of autonomous tissue dissection, which is the challenge of ensuring that all targeted tissue is fully severed, leaving no remaining connections. To address this challenge, an autonomous dissection pipeline shown in Fig. 2 is introduced, that has components of (1). autonomous dissection agents including both planning and learning-based methods (detailed in Sec III-A), (2). trajectory optimization for exposure maximization (Sec III-C), and (3). recovery planning that proposes future actions to achieve completeness (Sec III-D). Three components are designed to execute iteratively, providing actionable feedback to surgical dissection task.

A. Autonomous Dissection

We achieve autonomous dissection through two distinct paradigms: (1) a learned policy that executes bimanual retraction and cutting with visual inputs (2) a planning-based dissection agent that follows a 3D path generated by projecting tracked image keypoints into the robot's workspace.

3D Visuomotor Dissection Policy. Depth-based representations including point clouds from stereo matching and monocular depth are used as observations of the tissue. Because the state-of-the-art disparity estimation algorithm [23]

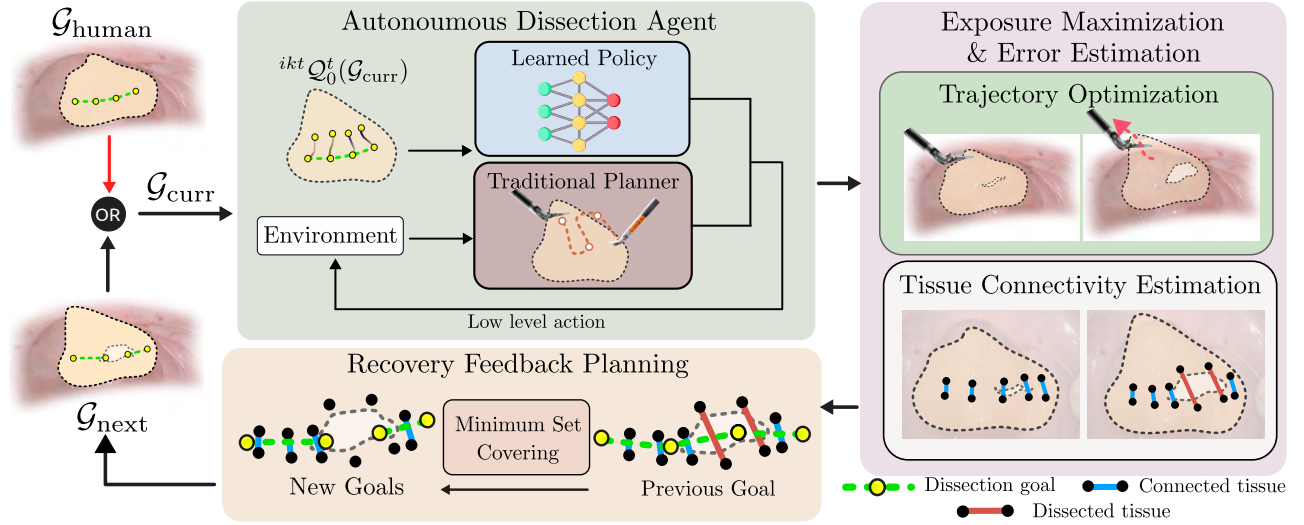


Fig. 2: **Our Autonomous Dissection Pipeline.** Our framework operates in a feedback loop. The process begins with a human-provided dissection goal $\mathcal{G}_{\text{human}}$, which the autonomous agent executes. Next, an exposure maximization controller manipulates the tissue to improve visibility for the subsequent error estimation. Finally, a recovery planner estimates the remaining tissue connectivity. If the dissection is incomplete, it generates a corrective goal, and the loop repeats until the task is successfully completed or a maximum number of attempts is reached.

prioritize accuracy over inference speed, we adopt a slow-fast strategy: high-frequency left monocular depth image, $\mathbf{M}_t^L \in \mathbb{R}^{3 \times H \times W}$, estimated by Video Depth Anything [24], and lower-frequency point clouds, $\mathbf{P}_t \in \mathbb{R}^{N \times 3}$, reconstructed from a modified Stereo Anywhere [23], in which Depth Anything v2 [25] is replaced with Video Depth Anything for monocular depth estimation in Stereo Anywhere for more consistent deep stereo matching.

We trained our policy via behavioral cloning on the human demonstrations we collected, using action chunking with transformer (ACT) [26] to generate a bi-manual action sequence given the current observations. The policy is learned to retract the tissue and perform dissection autonomously given a dissection goal from human demonstrations.

Planning-based Dissection. Given a tracked dissection goal \mathcal{G}^t and a grasping point $p^{\text{grasp}} \in \mathbb{Z}^2$ on the image plane initially defined from human experts, the planner generates a sequence of actions. First, it creates a 3D path to grasp the tissue and then retracts along the goal’s normal vector. The dissection trajectory is planned as a series of line segments, with the tool oriented along the path. Each cut is executed by moving to the target pose, inserting to a fixed depth, and closing the gripper to dissect.

B. Tissue Connectivity Estimation

We combine keypoint tracking with a rule-based approach to estimate tissue connectivity conditions undergoing tissue deformations (e.g. maximizing exposure covered in the next section). Defining the horizon of the estimation task as $t = t_d, \dots, t_e$, where t_d is the time immediately after a dissection is executed, and t_e denotes the end of the exposure and estimation procedure, the initial keypoint set $\mathcal{K}^{(t_d)} = \{\mathbf{k}_i^{(t_d)}\}_{i=1}^{N_k}$ is sampled as a uniform grid around the tracked dissection target \mathcal{G}^t . A temporary connectivity edge set $\mathcal{E} \subseteq \{e_{ij} \mid i, j \in \{1, 2, \dots, N_k\}, i \neq j\}$ is defined

base on the grid, including an undirected edge for every two neighboring grid points.

The tissue connectivity is estimated by iteratively computing the tracked keypoints temporal evolution,

$$\mathcal{K}^{(t_e)} = IKT \mathcal{Q}_{t_d}^{t_e}(\mathcal{K}^{(t_d)}), \quad (4)$$

and evaluating how much each edge’s length has changed. This elongation ratio is calculated for each edge as

$$\rho(e_{ij}) = \frac{\|\mathbf{k}_i^{(t_e)} - \mathbf{k}_j^{(t_e)}\|_2}{\|\mathbf{k}_i^{(t_d)} - \mathbf{k}_j^{(t_d)}\|_2}. \quad (5)$$

and an edge is viewed as disconnected (dissected) when it can be extended to τ times longer than its original length, i.e. $\rho(e_{ij}) \geq \tau$ where $\tau > 1$ is a ratio threshold. Finally, the connectivity set is updated to include only connected edges:

$$\mathcal{E}_t = \{e_{ij} \in \mathcal{E} \mid \rho_{ij} < \tau\}. \quad (6)$$

C. Exposure Maximization Control

The accuracy of our tissue connectivity estimation hinges on maximizing the elongation of tissue edges orthogonal to the dissection path. If this elongation is insufficient or poorly oriented, the geometric cues become ambiguous, making it difficult to reliably distinguish between connected and severed tissue as illustrated in Fig. 1. Therefore, two factors are critical for a robust tissue connectivity estimation: the magnitude of the elongation and its visibility to the camera. To address this, we propose an Exposure Maximization Controller that actively deforms the tissue to create an optimal viewing configuration to ensure robust tissue connectivity estimation. This controller solves an optimal control problem to find a robot action sequence that maximizes the visibility of the deformation.

Let \mathbf{M}_{t_d} be a surface mesh that is reconstructed from a segmentation and a point cloud [27] after a dissection action.

We do not make assumptions on the connectivity of the tissue as it is the estimation objective. Let $P_{t_d} \subset \mathbb{Z}^2$ be the set of image pixels near the proximity of a given dissection target. We create a simulation f_{xpbd} of \mathbf{M}_{t_d} , and formulate exposure maximization of P_{t_d} as the following finite horizon optimal control problem:

$$\begin{aligned} \min_{\mathbf{u}_{t_d:t_e}} \quad & \sum_{\tilde{t}=t_d}^{t_e} \left(\lambda_c \mathcal{L}_{\text{cam}}(P_{\tilde{t}}) + \lambda_d \mathcal{L}_{\text{deform}}(P_{\tilde{t}}) \right) \\ \text{subject to} \quad & \mathbf{x}_{\tilde{t}+1} = f_{\text{xpbd}}(\mathbf{x}_{\tilde{t}}, u_{\tilde{t}}), P_{\tilde{t}} = \text{mbt} \mathcal{Q}_{t_d}^{\tilde{t}}(P_{t_d}, \mathbf{x}_{t_d:\tilde{t}}), \\ & \mathbf{C}_{\text{vis}}(P_{\tilde{t}}) = 0, \tilde{t} = t_d, \dots, t_e \end{aligned} \quad (7)$$

where robot gripper position sequence $\mathbf{u}_{t_d:t_e} = \{u_{t_d}, \dots, u_{t_e}\}$ is initialized with the gripper position after dissection, $\mathcal{L}_{\text{cam}}, \mathcal{L}_{\text{deform}}$ and \mathbf{C}_{vis} are the costs and constraints to guide the action to maximize visibility. They are (1) camera-facing normal alignment loss to encourage surface normals to align with the viewing direction of the camera, (2) deformation expanding loss to encourage expansion of the dissected region, (3) a visibility constraint that enforces that $P_{\tilde{t}}$ are visible by the camera.

Camera Facing Loss: This loss penalizes deviations of the surface normals from the camera viewing direction. Let $\hat{P}_{\tilde{t}}$ be the corresponding 3D points on the simulated mesh from raycasting $P_{\tilde{t}}$, and let $N_{\tilde{t}}$ be corresponding surface normals. The loss is defined as

$$\mathcal{L}_{\text{cam}}(P_{\tilde{t}}) = \sum_{\mathbf{n}_i \in N_{\tilde{t}}} \mathbf{n}_i^\top \mathbf{w}_c, \quad (8)$$

where \mathbf{w}_c is the forward direction of the camera. The dot product $\mathbf{n}_i^\top \mathbf{w}_c$ is minimized when the vectors are antiparallel. Minimizing \mathcal{L}_{cam} encourages the dissection region to face toward the camera.

Deformation Expanding Loss: We hypothesize that tissue expansion in the direction that is orthogonal to the dissection goal will likely lead to ideal elongation of uncertain tissue connections, thus assists the proposed estimation module. To achieve this, we start by computing the deformation gradient matrix over the dissection area P_{t_d} , for the ease of decomposing the deformation into orthogonal and parallel components. The deformation gradient is computed as:

$$\mathbf{F}_{\tilde{t}}^p = \mathbf{I} + \nabla_p \mathbf{d}_{\tilde{t}}^p = \begin{bmatrix} 1 + \frac{\partial d_x^{(\tilde{t})}}{\partial x} & \frac{\partial d_x^{(\tilde{t})}}{\partial y} \\ \frac{\partial d_y^{(\tilde{t})}}{\partial x} & 1 + \frac{\partial d_y^{(\tilde{t})}}{\partial y} \end{bmatrix}, p \in P_{t_d} \quad (9)$$

where $\mathbf{d}_{\tilde{t}}^p$ is the 2D displacement of pixel p at time \tilde{t} w.r.t t_d . Geometrically, the deformation gradient $\mathbf{F}_{\tilde{t}}^p$ transforms a vector from its initial to its deformed configuration through a sequence of pure stretch and rigid-body rotation. Let $\tilde{\mathbf{n}}_{t_d}$ and $\tilde{\mathbf{m}}_{t_d}$ be unit vectors that are orthogonal and parallel to the dissection goal $\mathcal{G}^{(t_d)}$ at time t_d . Our goal here is to create a specific deformation behavior that expands the dissection region along $\tilde{\mathbf{n}}_{t_d}$ and reduces the region's shearing along $\tilde{\mathbf{m}}_{t_d}$. We achieve these objectives by minimizing the following function:

$$\mathcal{L}_{\text{deform}}(P_{\tilde{t}}) = \sum_{p \in P_{t_d}} -w_n (\tilde{\mathbf{n}}_{t_d}^\top \mathbf{F}_{\tilde{t}}^p \tilde{\mathbf{n}}_{t_d})^2 + w_s (\tilde{\mathbf{n}}_{t_d}^\top \mathbf{F}_{\tilde{t}}^p \tilde{\mathbf{m}}_{t_d})^2 \quad (10)$$

Algorithm 1 Greedy Next-Target Selection

- 1: **Input:** Uncut edges $\mathcal{E}^{\text{uncut}}$, segment length ℓ
- 2: **Output:** Next dissection target $\mathcal{G}_{\text{next}}$
- 3: Build candidate segments S
- 4: $U \leftarrow \mathcal{E}^{\text{uncut}}$
- 5: $\mathcal{G}_{\text{next}} \leftarrow \emptyset$
- 6: **while** $U \neq \emptyset$ **do**
- 7: $s^* \in \arg \max_{s \in S} \left| \{e \in U \mid s \cap e \neq \emptyset\} \right|$.
- 8: $\mathcal{G}_{\text{next}} \leftarrow \mathcal{G}_{\text{next}} \cup \{s^*\}$
- 9: $U \leftarrow U \setminus \{e \in U \mid s^* \cap e \neq \emptyset\}$
- 10: **end while**
- 11: **return** $\mathcal{G}_{\text{next}}$

where w_n and w_s are the hyperparameter we tune. $\tilde{\mathbf{n}}_{t_d}^\top \mathbf{F}_{\tilde{t}}^p \tilde{\mathbf{n}}_{t_d}$ quantifies how much the deformation keeps $\tilde{\mathbf{n}}_{t_d}$ aligned with itself and stretched along its own direction, and $\tilde{\mathbf{n}}_{t_d}^\top \mathbf{F}_{\tilde{t}}^p \tilde{\mathbf{m}}_{t_d}$ represents the loss of orthogonality under the deformation w.r.t the dissection goal.

We use a sampling-based optimal control solver for trajectory optimization. At each time step, we sample particles of future control trajectories with horizon h , simulating them in parallel and evaluating every particle with the proposed loss functions. To ensure visibility, we simply discard control sequences that leads to violation of the **visibility constraint**:

$$\mathbf{C}(P_{\tilde{t}+h}) = \begin{cases} 1, & \text{if } \exists (u, v) \in P_{\tilde{t}+h} \text{ s.t. } (u, v) \notin \Omega, \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

$$\Omega = \{(u, v) \in \mathbb{R}^2 \mid 0 \leq u < W, 0 \leq v < H\},$$

where W, H is the size of the endoscopic image. A nominal control trajectory $\hat{U} = \{\hat{u}_{\tilde{t}}, \dots, \hat{u}_{\tilde{t}+h}\}$ is obtained using the update rule from Model Predictive Path Integral (MPPI) [28], and we take a step forward in simulation with action $\hat{u}_{\tilde{t}+1}$. This procedure is implemented using GPU accelerated language Warp [29] to ensure fast optimization.

D. Recovery Feedback Planning

In this section, a method that detects changes in tissue connectivity as error metric of tissue dissection outcome is first discussed. It is then followed by a recovery module that proposes next dissection targets to the dissection autonomy to remove previous errors.

1) *Next Targets Proposal:* At the end of exposure maximization control ($t = t_d + t_e = T_{\text{end}}$), dissection along all segments in $\mathcal{G}^{(T_{\text{end}})}$ is successful if no remaining edge $e_{ij} \in \mathcal{E}_{T_{\text{end}}}$ intersects any target segment $g_m^{(T_{\text{end}})}$. The intersecting remaining edges form the uncut set

$$\mathcal{E}^{\text{uncut}} := \{e_{ij} \in \mathcal{E}_{T_{\text{end}}} \mid \exists m : e_{ij} \cap g_m^{(T_{\text{end}})} \neq \emptyset\}. \quad (12)$$

Assuming that the dissection policy/planner can reliably execute only a single cut segment of length at most ℓ per command, we seek a minimum-size set of commanded segments that removes all remaining edges. Let $\mathcal{G}_{\text{next}}$ be this set of command, the problem can be formulated as the

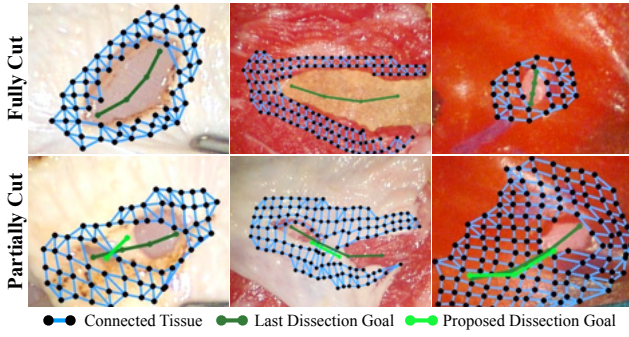


Fig. 3: Qualitative results of the recovery feedback planner across various tissues and conditions. After executing the last dissection goal, each sub-figure shows the resulting tissue connectivity and the planner’s next proposed target.

following minimum set covering problem,

$$\begin{aligned} & \min_{\mathcal{G}_{\text{next}}} |\mathcal{G}_{\text{next}}| \\ \text{s.t. } & \forall e \in \mathcal{E}^{\text{uncut}}, \exists s \in \mathcal{G}_{\text{next}} : s \cap e \neq \emptyset. \end{aligned} \quad (13)$$

This type of problem is known as NP-Hard, but can be solved using a greedy approximation algorithm as shown in Alg. 1. The first step of the algorithm is construction a set of command candidates by sampling around $\mathcal{E}^{\text{uncut}}$. Each candidate $s(c, \theta)$ is a line segment parametrized with center c , and rotation angle θ . The overall candidate set S is defined as

$$S = \{s(c, \theta) \mid \exists e_{ij} \in \mathcal{E}^{\text{uncut}} : d(e_{ij}, c) < \hat{d}, \theta \in [0, \pi]\}. \quad (14)$$

where $d(\cdot)$ is a distance function of a point to a line segments in 2D, and $\hat{d} > 0$ is a threshold parameter. In practice, position and orientation sampling is done by rejection sampling and bin sampling (bin size of 6), respectively.

After that, a greedy selector iteratively select the best candidate until $\mathcal{E}^{\text{uncut}}$ is covered. Let $U \leftarrow \mathcal{E}^{\text{uncut}}$ be the current uncover set, an element $e \in U$ is covered by a candidate s if they intersects (i.e. $s \cap e \neq \emptyset$). Candidates are ranked based on the number of element they can cover. The best candidate is draw as

$$s^* \in \arg \max_{s \in S} \left| \{e \in U \mid s \cap e \neq \emptyset\} \right|. \quad (15)$$

Then update the selection and uncovered set:

$$\mathcal{G}_{\text{next}} \leftarrow \mathcal{G}_{\text{next}} \cup \{s^*\}, U \leftarrow U \setminus \{e \in U \mid s^* \cap e \neq \emptyset\}. \quad (16)$$

Finally, $\mathcal{G}_{\text{next}}$ are given back to dissection controller as next dissection goals.

IV. EXPERIMENTS

We conduct extensive real-world experiments and detailed analysis to demonstrate the effectiveness of our method. All real-world experiments were performed on the da Vinci Research Kit(dVRK) [30]. Our algorithm controls two patient-side manipulators (PSM1, PSM2): one equipped with fenestrated bipolar forceps for tissue manipulation and retraction, and the other with pott scissors for dissection. To better

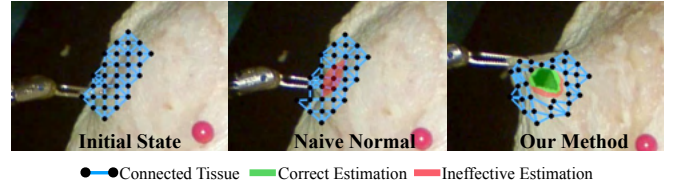


Fig. 4: Comparison of our exposure maximization control and naive normal for improving visibility of the dissected region. For the same amount of retraction, our method achieves significantly more exposure of the dissected area.

align with real clinical procedures for the real animal tissue experiment, we replaced the scissors with a cautery spatula tool.

A. Evaluation of Dissection Feedback

This section evaluates the two components of our dissection feedback framework. We first assess the performance of our tissue connectivity estimation across different image keypoint tracking methods. We then demonstrate the robustness and generalizability of our recovery feedback planning across a diverse set of tissues and conditions. We curated a dataset of 50 evaluation cases featuring a variety of materials: ex vivo animal tissues (feature-rich beef and feature-poor chicken skin) and silicone tissue phantoms. Each category includes both fully-cut and partially-cut samples to represent various states of dissection. To ensure a consistent and high-quality baseline, all exposure maximizing actions in the dataset were performed under expert human teleoperation.

To quantitatively evaluate the performance of our tissue connectivity estimation, we propose a metric, Detection Accuracy, defined as $\frac{N_{\text{correct}}}{N_{\text{total}}} \times 100\%$, where N_{total} is the total number of keypoint pairs evaluated, and N_{correct} is the number of pairs whose predicted separation state matches the ground-truth (connected or separated) labeled by human.

1) *Comparison of Different Tracking Methods:* CoTracker3 [31], RAFT [32], and LiteTracker [33] were evaluated for tissue connectivity estimation on our dataset. LiteTracker achieved the best average detection accuracy (92.8%), followed by CoTracker3 (90.5%), while RAFT performed substantially worse (73.4%). Given its highest accuracy and lowest computational overhead, we use LiteTracker in the remainder of this work.

2) *Quality of Recovery Feedback Planner:* We demonstrate the robustness and generalizability of our recovery feedback planner by testing it across diverse conditions, including different tissues (ex vivo beef and chicken skin), dissection states (partially and fully cut), and surgical instruments. The tools used were a mechanical cutter (Pott’s scissors) and an energy-based device (cautery spatula). Fig. 3 showcases the high quality of the dissection goals generated by our method in these scenarios.

B. Evaluation of Exposure Maximization Control

In this section, we evaluate the effectiveness of our exposure maximization control method. We conduct experiments on real animal tissue for 20 times to demonstrate that our

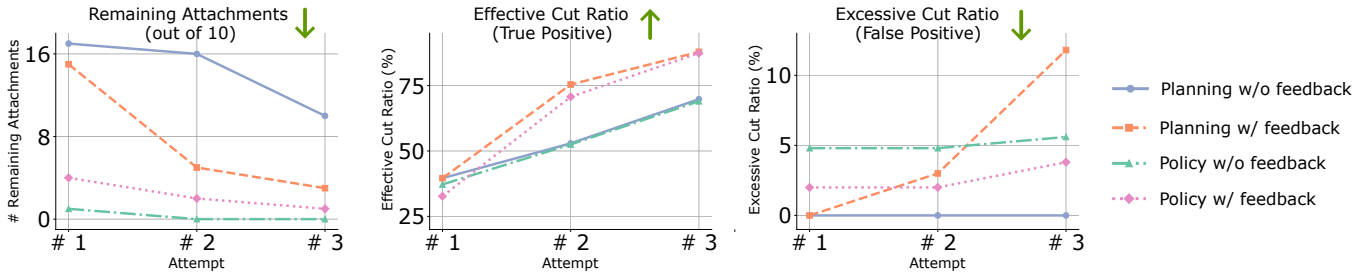


Fig. 5: Comparison of dissection performance over three attempts: (a) remaining attachments, (b) effective cut ratio, and (c) excessive cut ratio.

TABLE I: Comparison of visible area and connectivity estimation accuracy with and without our exposure maximization control.

	Visible Area (pxs)	Connect. Est. Acc. (%)
Ours-Normal	69.38 ± 60.68	5.56
Ours	1653.2 ± 506.57	86.00

control method not only enhances camera visibility but also improves the performance of the tissue connectivity estimation. For each trial, we place a real chicken tissue on a 20° slope with a predefined dissection goal and an initial dissected region on the chicken skin. We then employ our exposure maximization control to manipulate the tissue to maximize the camera visibility of this dissected region.

To quantify visibility, we manually contour the dissected region in the camera image and count the number of visible pixels on an 480 × 640 image. To assess the downstream benefit of improved visibility, we use a connectivity estimation accuracy defined as $\frac{N_{\text{pred}}}{N_{\text{GT}}} \times 100\%$. Here, N_{pred} is the number of edges predicted as dissected (i.e., separated), and N_{GT} is the ground-truth number of dissected edges labeled by a human expert.

Effects of Exposure Maximization Control. We analyze the effectiveness of our control method by comparing it against a naive normal planner. This baseline method plans a straight line trajectory in the direction of the normal vector computed from the grasping point and the midpoint of the dissection goal. For a fair comparison, both methods move a total of 0.5 mm per step for 35 steps. The results are summarized in Table I. With our method, the average visible area increased 23 times compared to the naive planner, indicating that our method successfully increases the exposure of the dissected area on real animal tissue. Furthermore, our method achieves a superior connectivity estimation accuracy. This is attributed to our controller’s ability to create more visible deformation around the dissected area, providing richer information to the tissue connectivity estimator, as illustrated in Fig. 4.

C. Evaluation of Dissection Autonomy with Feedback

In this section, we present experiments designed to demonstrate how our feedback framework improves the cutting performance of different autonomous dissection agents by enabling them to recover from errors. we constructed a scene with a volumetric silicon phantom containing a simulated

tumor, which is covered by a second, thin layer of textured phantom tissue. Our goal is to dissect the thin-shell.

We evaluate both autonomous dissection agents presented in Sec III-A for 60 times (each 10 times), highlighting the flexibility of our feedback approach. For each trial, an initial dissection goal is manually labeled on the image, and for the planning-based method, a grasping point is also manually specified.

As quantifying dissection performance in real-world scenarios is challenging, we assess the quality of each dissection using two primary metrics: (1) Length Deviation (mm): The difference between the length of the expected dissection path and the executed. (2) Remaining Attachments: The total number of uncut tissue strands remaining within the desired dissection region. Finally, we use these measurements to define a successful cut as Length Deviation of less than 2 mm and has zero Remaining Attachments.

Effects of dissection feedback. We compare three distinct conditions: (1) Ours (w/ Feedback): An autonomous dissection agent uses feedback to refine its dissection goal for up to three attempts. (2) Baseline (w/o Feedback, n=1): An agent without feedback that executes its dissection goal only once. (3) Baseline (w/o Feedback, n=3): An agent without feedback that executes the same initial dissection goal for three attempts. The second baseline was given three attempts to ensure a fair comparison by matching the maximum number of attempts of our feedback-enabled agents. Their results are summarized in Table II, showing that our feedback framework provides a significant performance improvement. With our dissection feedback framework, both autonomous agents achieve a much higher *Success Rate* than the baselines, demonstrating that our method effectively improves dissection performance across different dissection autonomy paradigms. Notably, the two agents exhibited distinct failure modes, showcasing the versatility of our feedback mechanism. The learning-based agent w/o feedback tended to perform clean cuts that were too short, failing to complete the entire goal. Our framework addressed this by identifying the uncut segment and generating a refined goal to complete the dissection. This recovery is quantitatively reflected in a significant drop in Length Deviation. In contrast, the planning-based agent was prone to leaving small tissue attachments, especially when the dissection path was long. Our framework successfully localized these uncut attachments and provided precise, corrective dissection goals to sever them. This is evidenced by a sharp decrease in the number of

TABLE II: Comparison of dissection performance on tissue phantoms. Here, n denotes the number of execution attempts.

Dissection Autonomy	Length Deviation (mm) ↓	Remaining Attachments (#) ↓	Success Rate (%) ↑
Planning-based w/o Feedback (n=1)	3.60 ± 1.44	16	20
Planning-based w/o Feedback (n=3)	3.40 ± 1.19	10	0
Learning-based w/o Feedback (n=1)	5.80 ± 2.27	0	45.5
Learning-based w/o Feedback (n=3)	4.19 ± 2.98	0	40
Planning-based w Feedback	1.47 ± 0.96	3	80
Learning-based w Feedback	1.84 ± 1.51	0	80

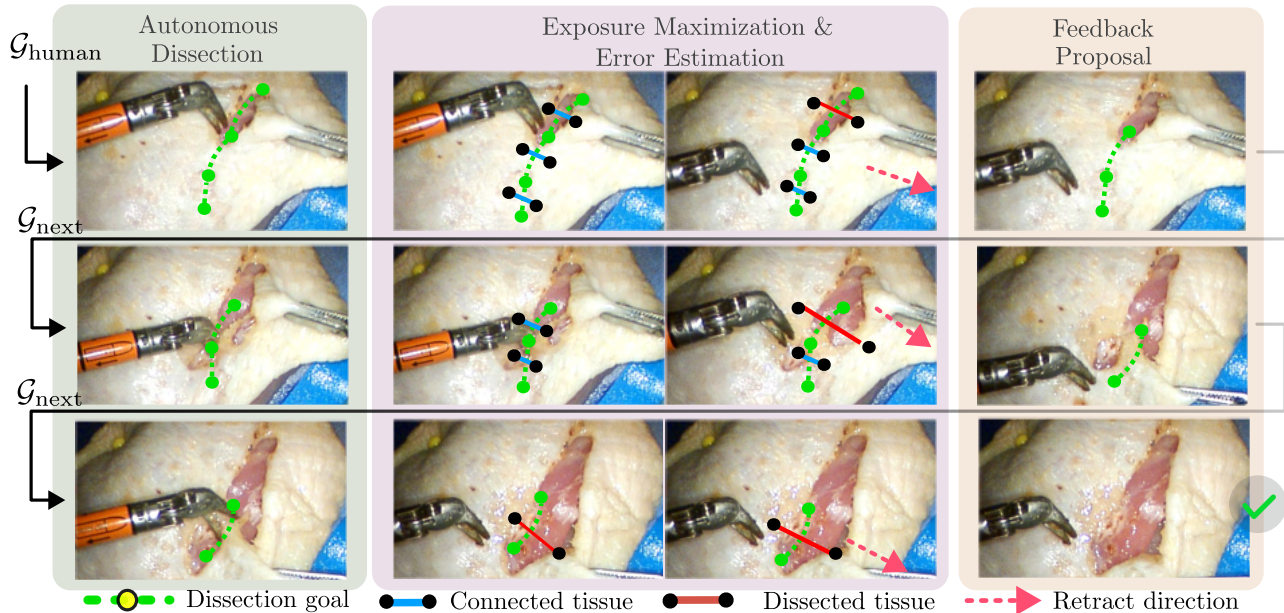


Fig. 6: This figure illustrates the execution of our complete feedback-enabled dissection pipeline. **(First Column)** The process begins with an agent executing a dissection goal. **(Second Column)** Next, our exposure maximization control retracts the tissue orthogonally to the cut to enhance visibility for error estimation. **(Third Column)** Then, our feedback planner identifies any remaining tissue attachments and generates a corrective goal. **(Bottom Right)** This loop repeats for up to three attempts until the dissection is complete.

Remaining Attachments. Without feedback, both paradigms performed poorly. While blindly re-executing the dissection three times offered a minor improvement over a single attempt, it was insufficient for achieving a clean, successful cut, highlighting the critical importance of our corrective feedback mechanism. To better understand error correction capabilities of our feedback mechanism, we evaluate the progression of dissection performance over three attempts, with results shown in Fig. 5. The results show that for the feedback-enabled agent, remaining attachments decrease sharply, the effective cut ratio increases significantly, and excessive cuts are suppressed over the three attempts. These trends demonstrate our feedback mechanism’s success in correcting imperfect dissections by simultaneously reducing different types of errors.

D. Real Tissue Demonstration

Finally, we demonstrate our complete framework (planning-based w feedback) on a long-horizon dissection task using ex vivo chicken tissue. Our framework successfully separate a 7.6 cm tissue segment sequentially, demonstrating its effectiveness in a realistic, multi-step scenario. Outcomes are shown in Fig. 6.

In the first column, our planning-based dissection agent execute a dissection given a dissection goal. Initially the goal was given by a human expert. Next our exposure maximization control deforms the tissue to maximize exposure for tissue connectivity estimation (second column). In the figure, the retraction direction is orthogonal to the dissection goal, making the deformation difference between the connected tissue and the dissected tissue more visible for better error estimation. Then, our recovery feedback planner provide a corrective goal as the next goal for the agent to execute. We limit the dissection attempts to 3 times to finally finish the dissection (bottom right).

V. DISCUSSION AND CONCLUSION

In this work, we introduce a feedback-enabled framework for autonomous tissue dissection. This framework is designed to reason about and localize topological changes following each dissection action, providing corrective, actionable feedback to the autonomous agent to recover from erroneous dissections. To ensure the robustness of this feedback, we propose a novel dissection visibility and an optimal controller that guides the robot to maximize this visibility through active tissue manipulation. Through extensive experiments

and detailed analysis of each component on both tissue phantoms and ex vivo animal tissues, we demonstrate the effectiveness of our method as a promising step toward robust dissection autonomy in surgery.

The primary limitation of our framework is its focus on thin-shell tissues, as it models and tracks keypoints only on the tissue surface. Consequently, the method may struggle to address occluded, partially folded, or volumetric tissues. However, recognizing that real-world surgical procedures predominantly require volumetric dissection, we seek to expand our feedback mechanism in future work to accommodate volumetric tissues. Furthermore, we intend to test the augmented framework on porcine tissue to establish its feasibility in realistic surgical settings.

REFERENCES

- [1] Z. Liang, Z.-Y. Chiu, F. Richter, and M. C. Yip, "Differentiable rendering-based pose estimation for surgical robotic instruments," 2025. [Online]. Available: <https://arxiv.org/abs/2503.05953>
- [2] C. D'Ambrosia, F. Richter, Z.-Y. Chiu, N. Shinde, F. Liu, H. I. Christensen, and M. C. Yip, "Robust surgical tool tracking with pixel-based probabilities for projected geometric primitives," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 15 455–15 462.
- [3] X. Liang, C.-P. Wang, N. U. Shinde, F. Liu, F. Richter, and M. Yip, "Medic: Autonomous surgical robotic assistance to maximizing exposure for dissection and cautery," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, 2025, pp. 6889–6895.
- [4] Z.-Y. Chiu, F. Richter, and M. C. Yip, "Real-time constrained 6d object-pose tracking of an in-hand suture needle for minimally invasive robotic surgery," *arXiv preprint arXiv:2210.11973*, 2022.
- [5] 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)*, 2024, pp. 15 471–15 477.
- [6] J. W. Kim, J.-T. Chen, P. Hansen, L. X. Shi, A. Goldenberg, S. Schmidgall, P. M. Scheickl, A. Deguet, B. M. White, D. R. Tsai *et al.*, "Srt-h: A hierarchical framework for autonomous surgery via language-conditioned imitation learning," *Science robotics*, vol. 10, no. 104, p. eadt5254, 2025.
- [7] N. U. Shinde, X. Liang, F. Liu, Y. Zhang, F. Richter, S. Herbert, and M. C. Yip, "Jiggle: An active sensing framework for boundary parameters estimation in deformable surgical environments," *Conference on Robotics: Science and Systems (RSS)*, 2024.
- [8] X. Liang, C.-P. Wang, N. U. Shinde, F. Liu, F. Richter, and M. Yip, "Medic: Autonomous surgical robotic assistance to maximizing exposure for dissection and cautery," 2024. [Online]. Available: <https://arxiv.org/abs/2409.14287>
- [9] F. Richter, S. Shen, F. Liu, J. Huang, E. K. Funk, R. K. Orosco, and M. C. Yip, "Autonomous robotic suction to clear the surgical field for hemostasis using image-based blood flow detection," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1383–1390, 2021.
- [10] A. Schüßler, C. Kunz, R. Younis, B. Alt, J. Paik, M. Wagner, and F. Mathis-Ullrich, "Semi-autonomous robotic assistance for gallbladder retraction in surgery," *IEEE Robotics and Automation Letters*, vol. 10, no. 7, pp. 7468–7475, 2025.
- [11] J. W. Kim, T. Z. Zhao, S. Schmidgall, A. Deguet, M. Kobilarov, C. Finn, and A. Krieger, "Surgical robot transformer (srt): Imitation learning for surgical tasks," *arXiv preprint arXiv:2407.12998*, 2024.
- [12] H. Saeidi, J. Ge, M. Kam, J. D. Opfermann, S. Leonard, A. S. Joshi, and A. Krieger, "Supervised autonomous electrosurgery via biocompatible near-infrared tissue tracking techniques," *IEEE Transactions on Medical Robotics and Bionics*, vol. 1, no. 4, pp. 228–236, 2019.
- [13] J. Ge, H. Saeidi, J. D. Opfermann, A. S. Joshi, and A. Krieger, "Landmark-guided deformable image registration for supervised autonomous robotic tumor resection," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, and A. Khan, Eds. Cham: Springer International Publishing, 2019, pp. 320–328.
- [14] K.-H. Oh, L. Borgioli, M. Zefran, L. Chen, and P. C. Giulianotti, "A framework for automated dissection along tissue boundary," *arXiv preprint arXiv:2310.09669*, 2023.
- [15] J. Ge, H. Saeidi, M. Kam, J. Opfermann, and A. Krieger, "Supervised autonomous electrosurgery for soft tissue resection," in *2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE)*. IEEE, 2021, pp. 1–7.
- [16] J. Ge, M. Kam, J. D. Opfermann, H. Saeidi, S. Leonard, L. J. Mady, M. J. Schnermann, and A. Krieger, "Autonomous system for tumor resection (astr) - dual-arm robotic midline partial glossectomy," *IEEE Robotics and Automation Letters*, vol. 9, no. 2, pp. 1166–1173, 2024.
- [17] Z. Karimi, S.-H. Ho, B. Thach, A. Kuntz, and D. S. Brown, "Reward learning from suboptimal demonstrations with applications in surgical electrocautery," *arXiv preprint arXiv:2404.07185*, 2024.
- [18] N. Marahrens, D. Jones, N. Murasovs, C. Biyani, and P. Valdastrì, "An ultrasound-guided system for autonomous marking of tumor boundaries during robot-assisted surgery," *IEEE Transactions on Medical Robotics and Bionics*, 2024.
- [19] T. François, L. Calvet, C. Sève-d'Erceville, N. Bourdel, and A. Bartoli, "Image-based incision detection for topological intraoperative 3d model update in augmented reality assisted laparoscopic surgery," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part IV 24*. Springer, 2021, pp. 647–656.
- [20] J. Ge, E. Kilmer, L. J. Mady, J. D. Opfermann, and A. Krieger, "Enhancing surgical precision in autonomous robotic incisions via physics-based tissue cutting simulation," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2024, pp. 2421–2428.
- [21] Y. Ou and M. Tavakoli, "Cressim-mpm: A material point method library for surgical soft body simulation with cutting and suturing," *arXiv preprint arXiv:2502.18437*, 2025.
- [22] N. Karaev, I. Rocco, B. Graham, N. Neverova, A. Vedaldi, and C. Ruppel, "Cotracker: It is better to track together," in *Proc. ECCV*, 2024.
- [23] L. Bartolomei, F. Tosi, M. Poggi, and S. Mattoccia, "Stereo anywhere: Robust zero-shot deep stereo matching even where either stereo or mono fail," in *Proceedings of the Computer Vision and Pattern Recognition Conference (CVPR)*, June 2025, pp. 1013–1027.
- [24] S. Chen, H. Guo, S. Zhu, F. Zhang, Z. Huang, J. Feng, and B. Kang, "Video depth anything: Consistent depth estimation for super-long videos," *arXiv:2501.12375*, 2025.
- [25] L. Yang, B. Kang, Z. Huang, Z. Zhao, X. Xu, J. Feng, and H. Zhao, "Depth anything v2," *arXiv:2406.09414*, 2024.
- [26] T. Z. Zhao, V. Kumar, S. Levine, and C. Finn, "Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardware," in *Proceedings of Robotics: Science and Systems*, Daegu, Republic of Korea, July 2023.
- [27] 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," *arXiv preprint arXiv:2309.11656*, 2023.
- [28] G. Williams, A. Aldrich, and E. Theodorou, "Model predictive path integral control using covariance variable importance sampling," *arXiv preprint arXiv:1509.01149*, 2015.
- [29] M. Macklin, "Warp: A high-performance python framework for gpu simulation and graphics," <https://github.com/nvidia/warp>, March 2022, nVIDIA GPU Technology Conference (GTC).
- [30] P. Kazanides, Z. Chen, A. Deguet, G. S. Fischer, R. H. Taylor, and S. P. DiMaio, "An open-source research kit for the da vinci® surgical system," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 6434–6439.
- [31] N. Karaev, I. Makarov, J. Wang, N. Neverova, A. Vedaldi, and C. Ruppel, "CoTracker3: Simpler and better point tracking by pseudo-labelling real videos," 2024.
- [32] Z. Teed and J. Deng, "Raft: Recurrent all-pairs field transforms for optical flow," 2020. [Online]. Available: <https://arxiv.org/abs/2003.12039>
- [33] M. A. Karaoglu, W. Ji, A. Abbas, N. Navab, B. Busam, and A. Ladikos, "Litetracker: Leveraging temporal causality for accurate low-latency tissue tracking," 2025. [Online]. Available: <https://arxiv.org/abs/2504.09904>