

Learning-Based Adaptive Control for Surgical Robotic Exposure Task on Deformable Tissues

Jiayi Liu, Kaiqi Wei, Yiwei Wang*, Huan Zhao and Han Ding
Huazhong University of Science and Technology, Wuhan, China

*Corresponding to: wang_yiwei@hust.edu.cn

Abstract—In various surgical procedures, regions of interest (ROIs) such as organs or lesions are often occluded by overlying tissues, requiring surgeons to achieve adequate exposure for precise intervention. However, the irregular geometry, nonlinear biomechanical properties of overlying tissues, and limited intraoperative visibility of the ROI pose significant challenges to the autonomous execution of tissue retraction. To address this, we formulate a realistic model of the tissue retraction task and propose a learning-based adaptive control framework for achieving ROI exposure. The method optimizes control inputs online by monitoring changes in the visual boundary of the tissue, while leveraging a deep deformation estimation model trained on simulation data to identify the optimal grasping point and ensure the convergence and safety of the adaptive controller. Through simulations and real-world experiments on different deformable materials, it has been demonstrated that this framework exhibits zero-shot adaptation to similar tasks and can complete the autonomous retraction process, from initial grasp selection to full ROI exposure. Therefore, it has the potential to be applied in actual surgical assistance scenarios.

I. INTRODUCTION

Robotic-assisted surgery (RAS) has achieved remarkable progress in clinical practice in recent years, significantly enhancing surgical outcomes across diverse procedures while effectively reducing the workload of medical personnel [3]. As research continues to advance the autonomy of surgical systems, an increasing number of auxiliary tasks demonstrate potential for independent execution by robotic platforms, thereby further liberating human surgical assistants [25]. Among these, tissue retraction is the most common auxiliary procedure in various surgeries [19]. For example, during a laparoscopic cholecystectomy, the gallbladder must be retracted to adequately expose the Calot’s triangle. Currently, such retraction remains heavily dependent on close coordination between surgical assistants and the primary surgeon. The autonomy of this operation will significantly reduce the physical and mental exertion of medical staff [2], but it faces some challenges: the irregular shape of occluding tissues impedes precise modeling, the highly nonlinear deformation behavior of biological tissues resists accurate prediction, and the region of interest (ROI) cannot be precisely determined before full exposure. These fundamental issues make autonomous tissue retraction one of the key focuses of current surgical robotics research.

To address these challenges, recent research in surgical robotics has increasingly focused on integrating advances in perception, modeling, and control to enable autonomous [14] or semi-autonomous [2, 27] tissue retraction. Early approaches

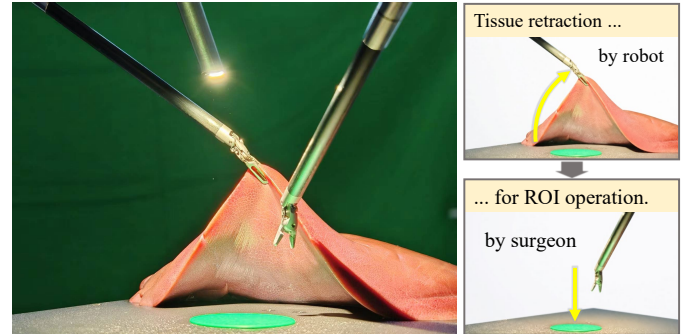


Fig. 1. The robot performs tissue retraction to expose the ROI, thereby enabling the surgeon to perform intervention on the obstructed area.

[19] relied on prior physics-based simulation models for optimization-based planning, which limited generalizability due to their strong dependence on these prior models. Subsequently, learning-based methods [4, 5, 21, 22, 26] that combine imitation learning and reinforcement learning have achieved end-to-end control, improving sample efficiency and safety through mechanisms such as contrastive representation and asymmetric input. Meanwhile, to enhance clinical interpretability, logic-based frameworks like FRAS [15] and DEFAS [28] explicitly encoded surgical domain knowledge using answer set programming, enabling failure-aware replanning and robust handling through online model updating. Moreover, DeformerNet [29, 30, 31] has been proposed as a point cloud-driven shape servoing framework that outputs end-effector displacement commands by learning the low-dimensional embedding difference between the current and target shapes, thereby enabling tissue retraction without requiring any physical prior model. Similarly, shape servoing forms the core technical foundation of this paper, as a general deformation control method.

Our method for shape servoing of deformable tissues revolves around the deformation Jacobian. This idea was first proposed by Navarro-Alarcón et al. [17, 18], who proposed a model-free adaptive shape control framework by estimating and updating the deformation Jacobian matrix of unknown elastic objects online. To enhance robustness, Lagneau et al. [11, 12] employed a sliding-window weighted least-squares method combined with an eigenvalue-based confidence criterion to update the Jacobian, effectively mitigating the impact of observation noise. Liu et al. [13] introduced a deep reinforcement learning agent to guide the adjustment of the Jacobian, improving convergence performance. Separately, to address

the problem of target point displacement caused by tissue deformation in surgical scenarios, Alambeigi et al. [1] and Zhong et al. [37] proposed estimating the deformation mapping under probe-tissue interaction to move the target point to a fixed instrument position in renal tumor cryoablation and suturing tasks, respectively. Recently, research has extended toward more sophisticated deformation representations. Yang et al. [34, 35] constructed low-dimensional, physically interpretable global deformation features based on modal analysis and derived an analytical form of the deformation Jacobian. Yu et al. [36] integrated offline pre-training with online adaptation, leveraging neural networks to learn a global Jacobian model for large-deformation control of deformable linear objects. Collectively, these advances demonstrate that adaptive control methods based on online deformation Jacobian estimation remain an effective paradigm for shape servoing tasks.

In this paper, we propose a unified framework that integrates learning and adaptive control for autonomous robotic tissue retraction tasks, with the explicit goal of exposing the ROI, as illustrated in Fig. 1. First, we design a set of low-dimensional features that effectively represent occlusion boundaries, and introduce the deformation Jacobian in this feature space to construct a model-free adaptive deformation control method. Then, we build a simulation for this retraction task to explore the interaction and deformation relationship between the manipulator and the tissue, and use this to train a data-driven deformation estimation model. The introduction of this model aims to improve the stability, reliability, and safety of the controller in actual operation. Importantly, we focus on the zero-shot transfer capability of the proposed method in real-world tasks, meaning that prior knowledge from simulation can be effectively integrated into the online adaptive adjustment process without retraining or fine-tuning in the real environment, thus achieving efficient and reliable retraction of unknown tissues.

The main contributions of our work are as follows:

- A novel tissue retraction framework is proposed. It estimates the deformation Jacobian by observing the tissue's boundary, and then uses it to drive the control system.
- A data-driven deep deformation estimation model is trained on simulation and is integrated to provide prior knowledge for the adaptive control pipeline.
- Real-world experiments are designed and conducted on ex vivo tissues to validate the effectiveness and practical application potential of our proposed method.

II. TASK FORMULATION

To model the tissue retraction task aimed at exposing an ROI while closely approximating the scenarios encountered in real surgical retraction procedures, this paper proposes the task formulation shown in Fig. 2. Specifically, a deformable elastic sheet is used to broadly represent the occluding tissue, with an irregularly shaped boundary to reflect the morphological diversity of real biological tissues. Additionally, a customizable two-dimensional elliptical marker with adjustable shape,

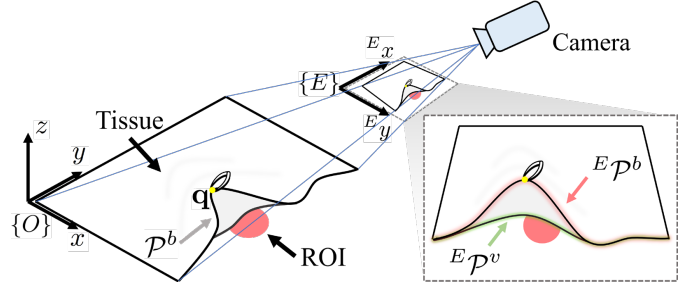


Fig. 2. Task formulation and key definitions for tissue retraction.

scale, position, and orientation is introduced to simulate the ROI that needs to be exposed in different surgical scenarios.

The retraction action $\mathbf{q}$ can be applied to any point along the tissue boundary $\mathcal{P}^b$ to produce deformation. Sensory feedback is derived from the 2D image provided by a laparoscopic camera that can be positioned at any location and orientation to capture the tissue. Here, ${}^E\mathcal{P}^b$ represents the projection of $\mathcal{P}^b$ onto the image plane, while ${}^E\mathcal{P}^v$ represents the visual boundary under the current actual occlusion state. The task is considered completed when ${}^E\mathcal{P}^v$ no longer overlaps with the projection of the ROI on the image plane, meaning the ROI is fully exposed and unobstructed in the field of view.

III. METHODOLOGY

This paper proposes a learning-based adaptive control method to accomplish the task of retracting occluding tissues for ROI exposure. Specifically, we first train a deformation estimation model in simulation. This model serves to initialize the deformation Jacobian and provide virtual deformation force constraints during the control iterations. Furthermore, we discuss the selection of the grasping position. Fig. 3 shows the overall framework of the proposed method.

A. Physics-Based Deformable Object Simulation

In order to collect prior deformation data, a simulation environment corresponding to the task formulation needs to be constructed. To balance accuracy and efficiency, the mass-spring model (MSM) [16, 23] is adopted for deformation computation. In this model, mesh vertices are treated as point masses, and edges between adjacent vertices are modeled as springs. The elastic force between two connected masses i and j follows Hooke's law:

$$\mathbf{F}_{ij} = k(\|\mathbf{p}_i - \mathbf{p}_j\| - l_0)\mathbf{e}_{ij}, \quad (1)$$

where k is the stiffness, l_0 is the rest length, and $\mathbf{e}_{ij}$ is the unit direction vector. The system is integrated using explicit Euler with timestep Δt and velocity damping $\zeta \in (0, 1]$:

$$\dot{\mathbf{p}} \leftarrow (\dot{\mathbf{p}} + \mathbf{F}/m_{eq} \cdot \Delta t) \cdot \zeta, \quad \mathbf{p} \leftarrow \mathbf{p} + \dot{\mathbf{p}} \cdot \Delta t, \quad (2)$$

where $\mathbf{p}$ and $\dot{\mathbf{p}}$ denote the positions and velocities, respectively, of the non-fixed masses. The deformable object's shape is updated iteratively using this integration rule, which is achieved through the hardware acceleration of Taichi [7].

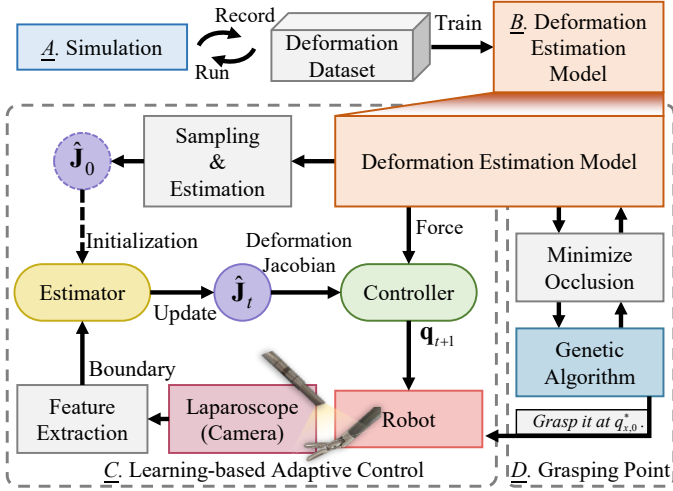


Fig. 3. Tissue retraction pipeline: The deformation estimation model is trained using data collected from simulations. Before the robotic retraction, it is employed to optimize the grasping point and initialize the deformation estimator. During retraction, it estimates deformation forces to serve as safety constraints for the controller. Tissue retraction is accomplished through continuous adjustment and execution within a closed-loop adaptive control.

Since the selection of the actual retraction strategies is influenced by the tissue boundary shape, the soft tissue model constructed for this simulation has randomized and irregular boundaries. The finite sine series $f(x) = \sum_{k=1}^K a_k \cdot \sin(k\pi x)$, $x \in [0, 1]$ is used to generate randomized boundaries, where $K \in \mathbb{Z}_{>0}$ is the number of harmonic terms controlling shape complexity, and $a_k \sim \mathcal{U}(-0.1, 0.1)$ are random amplitudes. This boundary profile function is scaled and applied to carve into a base cuboid grid of dimensions $100 \times 100 \times 4$, generating an occluding tissue model with a smoothly curved boundary.

In the task, the controlled variable is the position and orientation of the instrument, which is represented as $\mathbf{q} = [q_x, q_y, q_z, q_\alpha, q_\beta, q_\gamma]^T \in \mathbb{R}^6$. The initial grasping position $\mathbf{q}_0^p = [q_{x,0}, q_{y,0}, q_{z,0}]^T \in \mathcal{P}_0^b$ is chosen arbitrarily along the boundary, where $\mathcal{P}_0^b$ denotes the set of positions on the initial tissue boundary. The instrument begins at the initial pose $\mathbf{q}_0 = [(\mathbf{q}_0^p)^T, (\mathbf{0})^T]^T$. Grid nodes in the vicinity of the instrument's grasping point are rigidly displaced according to its motion, while the positions of all other non-fixed nodes are updated iteratively through the MSM.

B. Deep Deformation Estimation Model

During the preparation phase, we aim to train a deep neural network endowed with fundamental deformation knowledge, enabling it to guide the control process across a variety of similar retraction scenarios and compensate for deformation force information that is not readily measurable in the real world. The simulation environment is used to generate a dataset of action–deformation correspondences through random interaction, and the intermediate process of the MSM algorithm provides the relative values of the force vectors.

To represent the boundary $\mathcal{P}^b$, we employ an expansion in

terms of Wendland radial basis functions [33]:

$$\phi(\xi) = \sum_{j=1}^m w_j \varphi_{3,1} \left(\left| \frac{\xi - \xi_c^j}{h} \right| \right), \quad (3)$$

where ξ_c^j denotes the center of the j -th kernel, $h > 0$ is the support radius, and $\varphi_{3,1}$ is the Wendland kernel yielding a C^2 -continuous, compactly supported basis function. The number of kernels is denoted by $m \in \mathbb{Z}_{>0}$. Owing to their locality, smoothness, and numerical stability, Wendland kernels have been widely used in interpolation [6], function approximation [10], and soft tissue deformation modeling [32]. Importantly, both the derivatives and integrals of $\varphi_{3,1}$ admit closed-form expressions, which greatly simplify subsequent analytical computations. Using this kernel, we parameterize the 3D tissue boundary via a dedicated operator:

$$\mathbf{W}^b = \mathcal{R}^b(\mathcal{P}^b), \quad (4)$$

where $\mathcal{R}^b(\cdot)$ denotes the parameterization mapping that converts the tissue boundary $\mathcal{P}^b$ into a compact coefficient matrix $\mathbf{W}^b \in \mathbb{R}^{m \times 3}$. Full details of this construction, including kernel configuration, data centering, and least-squares fitting, are provided in Appendix I.

The maximum deformation force is of crucial importance for designing safety constraints. It is assumed to occur at grasped nodes $\mathcal{P}^g$. Since the MSM computes pairwise forces between mesh nodes during deformation simulation according to Equation (1), the resultant force on each grasped node can be calculated via summation over neighboring nodes, and the maximum deformation force can be extracted by:

$$F^{\max} = \max_{\mathbf{p}^i \in \mathcal{P}^g} \left\| \sum_{\mathbf{p}^j \in \mathcal{N}(i)} (\mathbf{F}_{ij}) \right\|_2, \quad (5)$$

where $\mathcal{N}(i)$ denotes the set of neighboring nodes of the i th node. The prediction model, while predicting deformation, also estimates the resulting deformation forces serving as a perception complement in scenarios without force sensors.

To achieve real-time prediction and estimation of shape and force, a multi-layer perceptron (MLP) with dimensions $(m+7)-64-256-512-256-(3m+1)$ is employed to construct the deep deformation estimation model:

$$[\text{vec}(\mathbf{W}^b)^T, F^{\max}]^T = \text{MLP}_\theta([\mathbf{w}_{x,0}^T, \mathbf{q}^T, q_{x,0}]^T), \quad (6)$$

where $\text{vec}(\mathbf{W}^b)$ represents the vectorization of $\mathbf{W}^b$ with $3m$ parameters. $\mathbf{w}_{x,0} \in \mathbb{R}^m$ denotes the column vector of $\mathbf{W}_0^b = [\mathbf{w}_{x,0}, \mathbf{0}, \mathbf{0}]$ corresponding to the x-coordinate component. Since the initial boundary point set $\mathcal{P}_0^b$ lies on the x-y plane, the boundary shape can be parameterized by x along the y-direction, reducing the number of required parameters. After reconstructing the boundary curve, the grasping position constrained to lie on the curve can be uniquely determined by specifying $q_{x,0}$ alone.

C. Learning-Based Adaptive Control

The learning-based adaptive control method leverages the trained deformation model to guide actions and enforce safety

constraints. The retraction task in this phase is defined in a 2D coordinate frame $\{E\}$, which lies in the plane of the camera image. The x-axis of $\{E\}$ is aligned with the line connecting the start and end points of the observed tissue boundary ${}^E\mathcal{P}^b$, while the y-axis is oriented parallel to the camera's vertical axis. The origin of $\{E\}$ is placed at the starting point of the observed tissue boundary. ${}^E\mathcal{P}^b$ is defined as the projection of the tissue boundary onto the image plane. In the early stage of retraction, it is almost consistent with ${}^E\mathcal{P}^v$.

To achieve the exposure task in closed-loop control using only camera image feedback, we focus on the visual boundary ${}^E\mathcal{P}^v$. Due to differences in dimensionality and in the sampling strategy used to extract coordinates from the simulation, we employ a separate parameterization operator:

$$\mathbf{w}^v = \mathcal{R}^v({}^E\mathcal{P}^v), \quad (7)$$

where $\mathcal{R}^v(\cdot)$ denotes the parameterization with full details shown in Appendix II. The resulting parameter vector $\mathbf{w}^v \in \mathbb{R}^m$ encodes the shape of ${}^E\mathcal{P}^v$.

The following part presents the modeling of the retraction task. It is essential to clarify that this task is exploratory. Before the ROI is exposed, retraction decisions can only be made based on the currently visible regions. The primary objective is to reduce the occlusion area of the overlapping parts between the tissue and the ROI. The loss function reflecting the occlusion is designed as follows:

$$\begin{aligned} \mathcal{L} &= \int_{\mathcal{D}} \phi^v({}^Ex) d{}^Ex \\ &= \int_{\mathcal{D}} \sum_{i=1}^m {}^i\mathbf{w}^v \varphi \left(\left| \frac{{}^Ex - {}^Ex_c^i}{h} \right| \right) d{}^Ex, \end{aligned} \quad (8)$$

where $\mathcal{D} = \cup_s [x_{Ls}, x_{Hs}]$ represents the overlap area between the boundary of the observable ROI and that of the upper tissue. The overlap area is generally a single segment, that is, $s = 1$. There could also be multiple segments, that is, $s > 1$. Furthermore, if the ROI is fully covered, the integration area is set to the entire domain.

The control objective is to minimize the loss function in the direction of the optimal action, achieving efficient exposure. The partial derivative of $\mathcal{L}$ with respect to $\mathbf{q}$ is computed to obtain the pose update for the instrument:

$$\frac{d}{dt} \mathbf{q} = -\kappa \frac{\partial \mathcal{L}}{\partial \mathbf{q}} = -\kappa \left(\frac{\partial \mathbf{w}^v}{\partial \mathbf{q}} \right)^T \frac{\partial \mathcal{L}}{\partial \mathbf{w}^v}, \quad (9)$$

where κ is a positive constant value. The second term captures the relationship between the occluded area and the visual boundary parameters, and can be computed by combining it with Equation (8):

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}^v} = \left[\int_{\mathcal{D}} \varphi \left(\left| \frac{{}^Ex - {}^Ex_c^i}{h} \right| \right) d{}^Ex \right]_{i=1}^m. \quad (10)$$

An adaptive deformation Jacobian $\mathbf{J} \in \mathbb{R}^{m \times 6}$ is introduced to describe the relationship between visual boundary parameters

and the instrument's actions. Based on $\mathbf{J}$, the first term in Equation (9) can be determined as:

$$\frac{\partial \mathbf{w}^v}{\partial \mathbf{q}} = \mathbf{J}. \quad (11)$$

To obtain an accurate real-time Jacobian estimate, it must be reasonably initialized during the setup phase to ensure convergence. Subsequently, during control iterations, the Jacobian is adaptively adjusted based on the observed estimation error. The deformation estimation model is used to participate in the estimation of $\mathbf{J}_0$. First, ν sets of random actions are generated $\mathbf{Q}_0 = [\delta \mathbf{q}_0^1, \delta \mathbf{q}_0^2, \dots, \delta \mathbf{q}_0^\nu] \in \mathbb{R}^{6 \times \nu}$. The model infers the corresponding boundary parameters for each action based on the grasping position and the initial boundary parameters. The boundary deformation parameters estimated by the model are as follows:

$${}^E\hat{\mathbf{W}}_0^b = [\delta {}^E\mathbf{w}_0^{b,1}, \dots, \delta {}^E\mathbf{w}_0^{b,\nu}] \in \mathbb{R}^{m \times \nu}, \quad (12)$$

$$\delta {}^E\mathbf{w}_0^{b,i} = {}^E\mathcal{T}(\hat{\mathbf{W}}_0^{b,i}) - {}^E\mathbf{w}_0^b, \quad i \in \{1, \dots, \nu\}, \quad (13)$$

where $\hat{\mathbf{W}}_0^{b,i}$ represents the boundary parameters being estimated by the model in the $\{O\}$ coordinate, and ${}^E\mathcal{T}(\cdot)$ is the function that transforms it into the $\{E\}$ coordinate. Since around the initial state, the visual boundary ${}^E\mathcal{P}^v$ is close to the tissue boundary ${}^E\mathcal{P}^b$, it can be approximately considered that ${}^E\hat{\mathbf{W}}_0^v \approx {}^E\hat{\mathbf{W}}_0^b$. Then the initial Jacobian can be estimated by the relationship $\langle {}^E\hat{\mathbf{W}}_0^b, \mathbf{Q}_0 \rangle$:

$${}^E\hat{\mathbf{W}}_0^b = \hat{\mathbf{J}}_0 \mathbf{Q}_0. \quad (14)$$

This can be formulated as a least-squares problem:

$$\underset{\hat{\mathbf{J}}_0}{\text{minimize}} \|\mathbf{Q}_0^T \hat{\mathbf{J}}_0^T - ({}^E\hat{\mathbf{W}}_0^b)^T\|_F^2. \quad (15)$$

Its analytical solution is given by:

$$\hat{\mathbf{J}}_0 = {}^E\hat{\mathbf{W}}_0^b \mathbf{Q}_0^T (\mathbf{Q}_0 \mathbf{Q}_0^T)^{-1}, \quad (16)$$

where $\mathbf{Q}_0 \mathbf{Q}_0^T$ is a nonsingular matrix because it is from random sampling, hence, the quadratic form can be inverted.

During manipulation, the deformation Jacobian changes, so the estimated values need to be adjusted. During each control iteration, the parameter changes resulting from the execution of the action can be estimated as follows:

$$\delta \hat{\mathbf{w}}_{t-1}^v = \hat{\mathbf{J}}_{t-1} \delta \mathbf{q}_{t-1}. \quad (17)$$

The actual visual boundary parameters can be observed through the image. The error between the theoretical estimation and the actual parameters is denoted by $\mathbf{e}$:

$$\mathbf{e} = \hat{\mathbf{J}}_{t-1} \delta \mathbf{q}_{t-1} - \delta \mathbf{w}_{t-1}^v \in \mathbb{R}^m. \quad (18)$$

To maintain the estimated $\hat{\mathbf{J}}$ close to the deformation model, the algorithm adapts to match the estimated error. It employs a gradient descent method to minimize a quadratic function:

$$\mathcal{H} = \mathbf{e}^T \mathbf{e} / 2. \quad (19)$$

Update $\hat{\mathbf{J}}$ according to the gradient rule:

$$\frac{d}{dt} {}^i\hat{\mathbf{J}} = -\lambda \frac{\partial \mathcal{H}}{\partial {}^i\hat{\mathbf{J}}} = -\lambda {}^i\mathbf{e} \delta \mathbf{q}_{t-1}^T, \quad (20)$$

where $\lambda > 0$ is the adaptation rate. ${}^i\hat{\mathbf{J}}$ represents the row vector of $\hat{\mathbf{J}}$, and ${}^i\mathbf{e}$ denotes the element of $\mathbf{e}$. In the discrete control iterations, the update rule takes the form of:

$$\hat{\mathbf{J}}_t = \hat{\mathbf{J}}_{t-1} - \lambda \mathbf{e} \delta \mathbf{q}_{t-1}^T. \quad (21)$$

In practical retractions, excessive deformation forces may cause tissue damage. Therefore, force constraints are imposed on instrument manipulation. Building upon the loss function in Equation (8), a force penalty term is added:

$$\mathcal{L}^* = \mathcal{L} + \tau \mathcal{L}^F \left(\hat{F}^{\max} - F^{\text{safe}} \right), \quad (22)$$

where τ denotes the weight of the force penalty. $\hat{F}^{\max}$ is the maximum deformation force estimated by the network for $\mathbf{q}_t$. F^{safe} represents the maximum force threshold that the tissue can safely withstand. The form of $\mathcal{L}^F(x)$ is as follows:

$$\mathcal{L}^F(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ x^2, & \text{if } x > 0. \end{cases} \quad (23)$$

It can be seen that when the estimated force is less than the safety value, the additional term is 0, so the action update is consistent with Equation (9). However, when the estimated force exceeds the safety value, the update rule becomes:

$$\frac{d}{dt} \mathbf{q} = -\kappa_1 \frac{\partial \mathcal{L}}{\partial \mathbf{q}} - \kappa_2 \frac{\partial \mathcal{L}^F}{\partial \mathbf{q}}, \quad (24)$$

where $\kappa_2 = \tau \kappa_1$. The method for solving the first partial derivative is the same as that mentioned earlier. The partial derivative of the penalty term is solved as follows:

$$\begin{aligned} \frac{\partial \mathcal{L}^F}{\partial \mathbf{q}} &= \frac{\partial \mathcal{L}^F}{\partial \hat{F}^{\max}} \cdot \frac{\partial \hat{F}^{\max}}{\partial \mathbf{q}} \\ &= 2 \left(\hat{F}^{\max} - F^{\text{safe}} \right) \cdot \text{MLP}_{\theta}^{-1} \left(\hat{F}^{\max} \right), \end{aligned} \quad (25)$$

where $\text{MLP}_{\theta}^{-1}(\cdot)$ represents the backpropagation process, which enables the calculation of the gradient relationship between the input and output of the neural network.

To form the controller, $\mathbf{u}_t = d\mathbf{q}/dt = [(\mathbf{u}_t^v)^T, (\mathbf{u}_t^{\omega})^T]^T$ is used to represent the desired increment of the instrument's position. The maximum speeds for both position and orientation need to be restricted separately:

$$\mathbf{u}_t^{v/\omega,*} = \min \left(\mathbf{u}_t^{v/\omega}, s^{v/\omega} \cdot \mathbf{u}_t^{v/\omega} / \|\mathbf{u}_t^{v/\omega}\|_2 \right), \quad (26)$$

where scalar $s^{v/\omega}$ determines the maximum linear or angular velocity during the robot's manipulation. The function $\min(\cdot)$ returns the vector with the smaller modulus from the two vectors provided. The final controller is as follows:

$$\mathbf{q}_{t+1} = \mathbf{q}_t + [(\mathbf{u}_t^{v,*})^T, (\mathbf{u}_t^{\omega,*})^T]^T. \quad (27)$$

The robot will control the grasper according to this rule until the ROI is fully exposed ($\mathcal{D} = \emptyset$).

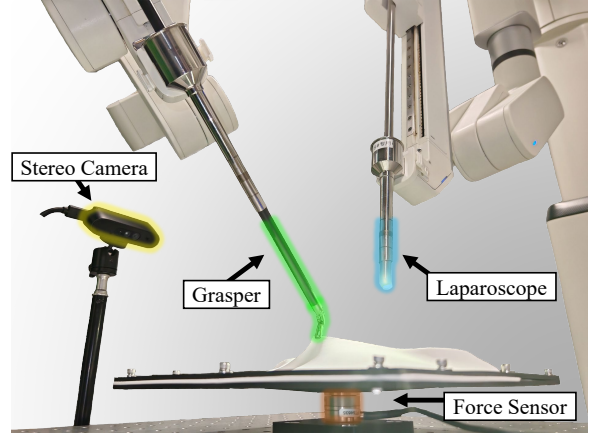


Fig. 4. Real-world experimental setup.

D. Preoperative Grasping Point Optimization

Different grasping positions significantly affect the retraction action and exposure efficiency, and once selected, the grasping position cannot be changed during the continuous retraction process. Therefore, choosing an appropriate grasping position is another key issue that needs to be addressed.

To obtain the optimal grasping position under partial observations, the idea is to find the position where the action performed maximizes the exposure of the initial observed ROI. Additionally, the deformation produced by this action must be within the safe range. An optimization problem with inequality constraints is formulated as follows:

$$\underset{q_{x,0}, \mathbf{q}}{\text{minimize}} \left(\hat{\mathbf{w}}^v(q_{x,0}, \mathbf{q}) - {}^E \mathbf{w}_0^b \right)^T \left(\epsilon \frac{\partial \mathcal{L}_{\mathcal{D}}}{\partial \mathbf{w}^v} + (1 - \epsilon) \frac{\partial \mathcal{L}_{\Omega}}{\partial \mathbf{w}^v} \right), \quad (28)$$

$$\text{s.t. } \hat{F}^{\max}(q_{x,0}, \mathbf{q}) < F^{\text{safe}}, \quad (29)$$

where $\mathcal{L}_{\mathcal{D}}$ and $\mathcal{L}_{\Omega}$ respectively represent the shaded area on the boundary of coverage and the shaded area on the entire boundary. The hyperparameter ϵ is used to balance the two aspects. A larger ϵ emphasizes the observed ROI exposure, while a smaller value emphasizes the overall exposure. In the constraint, the deformation estimation model takes in the action and provides an estimated deformation force, which is limited to be less than the safe force. This optimization demonstrates the process of selecting the action that achieves the best exposure with a limited retracting force.

To obtain the optimal $q_{x,0}^*$, the inequality constraints are first implicitly incorporated into the objective function:

$$I_{-}(q_{x,0}, \mathbf{q}) = \begin{cases} +\infty, & F^{\Delta} \geq 0, \\ -(1/t) \log(-F^{\Delta}), & F^{\Delta} < 0, \end{cases} \quad (30)$$

where $F^{\Delta} = \hat{F}^{\max}(q_{x,0}, \mathbf{q}) - F^{\text{safe}}$, and the approximate optimization function is:

$$\begin{aligned} \underset{q_{x,0}, \mathbf{q}}{\text{minimize}} \left(\hat{\mathbf{w}}^v(q_{x,0}, \mathbf{q}) - {}^E \mathbf{w}_0^b \right)^T \left(\epsilon \frac{\partial \mathcal{L}_{\mathcal{D}}}{\partial \mathbf{w}^v} + (1 - \epsilon) \frac{\partial \mathcal{L}_{\Omega}}{\partial \mathbf{w}^v} \right) \\ + I_{-}(q_{x,0}, \mathbf{q}). \end{aligned} \quad (31)$$

Then, the genetic algorithm is employed to find the optimal solution $q_{x,0}^*$ for this optimization problem. The grasping

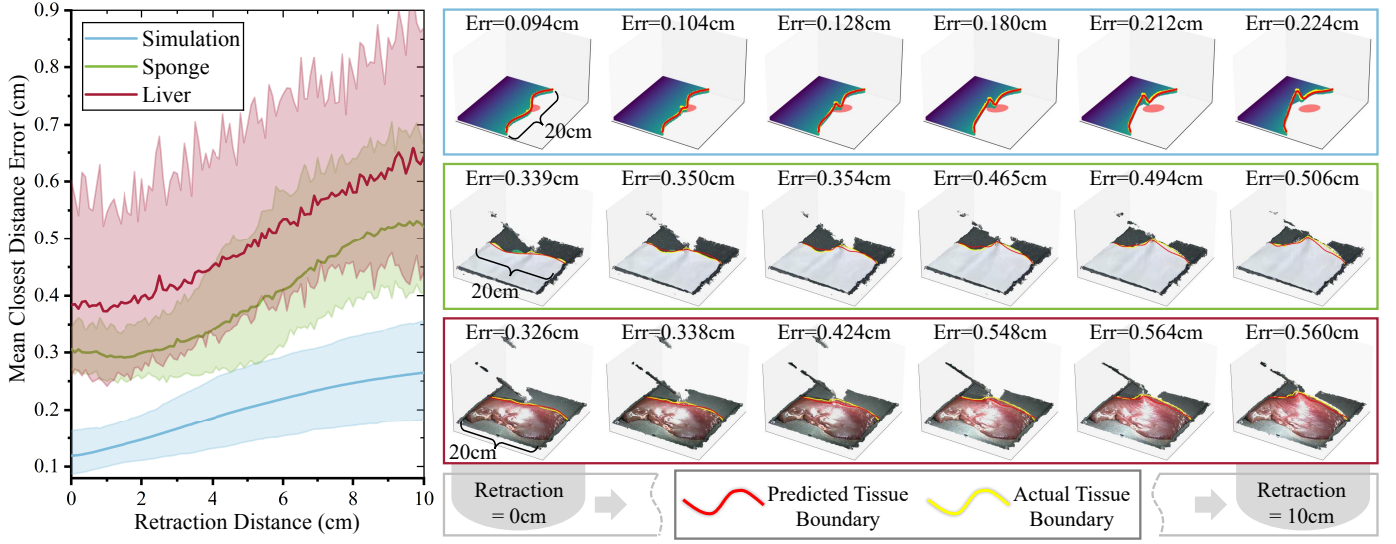


Fig. 5. (Left) The mean closest distance error, which is computed between the deformation network’s predicted boundaries and the measured boundaries in relation to the retraction distance across the three experimental scenarios. For each retraction distance, the solid line denotes the mean error, and the translucent band represents the 10th to 90th percentile range. (Right) A retraction process is visualized for each experimental scenario, with the predicted tissue boundary shown in red and the actual tissue boundary in yellow.

position is selected during the offline phase, and it needs to be recalculated for different boundaries and ROIs. After the grasping position is determined, the grasper is bound to the tissue, and subsequent actions are decided by the controller.

The complete process of the learning-based adaptive control method for exposure tasks is summarized in Algorithm 1.

Algorithm 1 Learning-Based Adaptive Control

Input: Visual boundary parameters $\mathbf{w}_t^v$

Output: Control policy (robot’s next state) $\mathbf{q}_{t+1}$

- 1: Train deformation estimation model
- 2: Select initial grasping pose $\mathbf{q}_0$ via genetic algorithm
- 3: Sample around $\mathbf{q}_0$ to obtain $\mathbf{Q}_0$
- 4: Estimate deformation ${}^E\hat{\mathbf{W}}_0^b$
- 5: Compute initial Jacobian $\hat{\mathbf{J}}_0 = {}^E\hat{\mathbf{W}}_0^b \mathbf{Q}_0^T (\mathbf{Q}_0 \mathbf{Q}_0^T)^{-1}$
- 6: **while** $\mathcal{D} \neq \emptyset$ (until exposure) **do**
- 7: Observe current visual boundary $\mathbf{w}_t^v$
- 8: Estimate error $\mathbf{e} = \hat{\mathbf{J}}_{t-1} \delta \mathbf{q}_{t-1} - \delta \mathbf{w}_{t-1}^v$
- 9: Update Jacobian $\hat{\mathbf{J}}_t = \hat{\mathbf{J}}_{t-1} - \lambda \mathbf{e} \delta \mathbf{q}_{t-1}^T$
- 10: Compute velocity $\mathbf{u}_t = \frac{d}{dt} \mathbf{q} = -\kappa_1 \frac{\partial \mathcal{L}}{\partial \mathbf{q}} - \kappa_2 \frac{\partial \mathcal{F}}{\partial \mathbf{q}}$
- 11: Limit velocity $\mathbf{u}_t^* = \min(\mathbf{u}_t, s \cdot \mathbf{u}_t / \|\mathbf{u}_t\|_2)$
- 12: Update controller $\mathbf{q}_{t+1} = \mathbf{q}_t + \mathbf{u}_t^*$
- 13: Command grasper to reach $\mathbf{q}_{t+1}$
- 14: **end while**

IV. EXPERIMENTS

A. Experimental Setup

To validate the performance of the proposed method, three experimental scenarios were designed. First, simulations were conducted in a virtual environment, where tissue deformation parameters remained identical to those used during training data acquisition for the deformation estimation network, while boundary geometries and ROIs were randomized. To align

with real-world dimensions, the dimension of the simulated tissue was scaled to 20 cm.

Subsequently, real-world experiments employed two deformable specimens: polyvinyl alcohol (PVA) sponge [8] and ex vivo porcine liver tissue. Compared to simulation, these materials exhibit distinct deformation characteristics, with the porcine liver further demonstrating complex geometry and tissue heterogeneity [20]. All real-world experiments utilized the MP550 surgical system from United Imaging Surgical. One robotic arm mounted an endoscopic camera, prepositioned manually to ensure boundary visibility. The other arm was equipped with a grasper to clamp tissue boundaries and execute retraction under the control of the proposed algorithm.

To capture ground-truth deformation data, an Orbbec Gemini 336 stereo camera and an ATI Nano25 SI-250-6 force sensor were deployed. The camera reconstructed the surface point cloud and identified boundary points through color thresholding segmentation for real-time boundary estimation. The force sensor was installed beneath the support platform and was used to measure the force applied to the tissue. The retraction force F^r was derived from the principle of force equilibrium: $F^r \approx \|\mathbf{F}^{\text{sensor}}\|_2$, where $\mathbf{F}^{\text{sensor}} \in \mathbb{R}^3$ represents the three-dimensional force measured by the sensor. The real-world experimental setup is illustrated in Fig. 4.

B. Deformation Estimation

Before performing autonomous tissue manipulation, a deformation estimation network was trained and experimentally validated. The training data were collected as follows: the simulated tissue boundary was randomized 100 times; for each randomized shape, 10 grasping points were uniformly sampled along the boundary. For each shape-grasping-point pair, a 10-cm pulling action was applied in a randomly sampled direction. To ensure physical plausibility, the sampling range of the direction was constrained to the collision-free

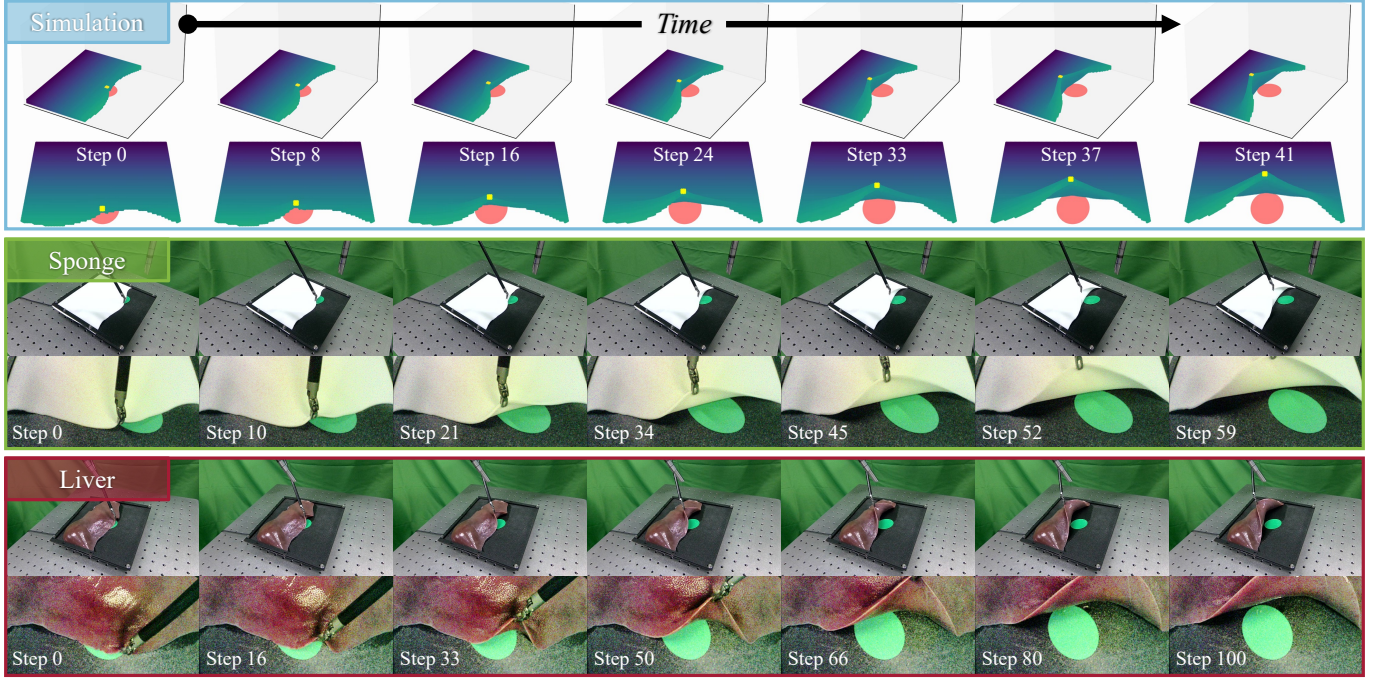


Fig. 6. Snapshots from a representative autonomous retraction trial on simulation, PVA sponge, and porcine liver. Each set shows two views: the upper row displays the global perspective captured by an external camera, while the lower row presents the laparoscopic perception view.

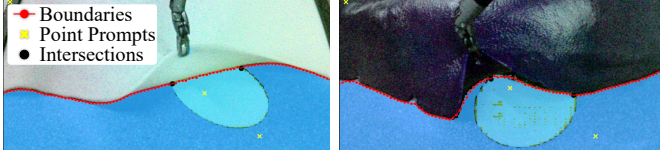


Fig. 7. SAM2 was chosen to obtain the visual boundaries of the tissue.

workspace. During each retraction, 100 time steps of data were recorded. This procedure yielded 10^5 structured samples, each comprising the corresponding tissue boundary parameters, maximum deformation force, grasping location, and action.

The deep deformation estimation network was trained on the aforementioned simulated dataset. The loss function was defined as the mean squared error (MSE) between the predicted and ground-truth tissue deformation values. The MLP architecture employed LeakyReLU activation functions in all hidden layers and a linear activation in the output layer. Training was performed using the Adam optimizer [9] with a batch size of 32, an initial learning rate of 1×10^{-3} , and default momentum parameters ($\beta_1 = 0.9$, $\beta_2 = 0.999$). The learning rate was dynamically halved via a ReduceLROnPlateau scheduler when the validation loss plateaued. Convergence was achieved after approximately 400 epochs. The dataset was partitioned into training and validation sets in a 9:1 ratio, and additional online evaluation was conducted within the simulation environment to assess real-time prediction performance.

To validate the performance of the deep deformation estimation network, a unified operational protocol was executed across three experimental scenarios: for tissues with varying geometric morphologies, retraction with a total displacement of 10 cm was applied from multiple grasping positions along a random direction. In each trial, the ground-truth tissue boundary was reconstructed from mesh or point cloud data

and compared against the predicted boundary, which was generated by reconstructing the network's output parameters. The mean closest distance (MCD) was adopted as the metric for prediction error, defined as:

$$\text{MCD}(\mathcal{P}^b, \hat{\mathcal{P}}^b) = \frac{1}{N_{\mathcal{P}^b}} \sum_{\mathbf{p}^b \in \mathcal{P}^b} \min_{\hat{\mathbf{p}}^b \in \hat{\mathcal{P}}^b} \|\mathbf{p}^b - \hat{\mathbf{p}}^b\|_2, \quad (32)$$

where $\hat{\mathcal{P}}^b$ denotes the predicted boundary, and $N_{\mathcal{P}^b}$ is the number of sampled points on the actual boundary. Since the predicted boundary is continuous and can be sampled at far more points ($\hat{\mathbf{p}}^b$) than the measured points ($\mathbf{p}^b$), only the unidirectional MCD error is computed. A total of 1,000 simulations and 100 real-world experiments were conducted. Fig. 5 presents the error distribution across all experiments and visualizes a retraction process from each scenario.

The results demonstrate that the model achieves low prediction errors in simulated scenarios that align with the training data distribution. Although the error increases on real tissues (including sponge and liver), it remains within a clinically acceptable range (< 1 cm for retractions up to 10 cm), and the model accurately captures the overall deformation trend. Moreover, the model exhibits lower error under small displacement conditions, effectively supporting the adaptive control method, which relies on local sampling and inference around the initial configuration. In summary, this study provides preliminary validation that the proposed deep deformation estimation network can reliably support critical components of both simulation and real-world tasks.

C. Adaptive Retraction

The proposed autonomous tissue retraction method was evaluated in three experimental scenarios. In simulation, the

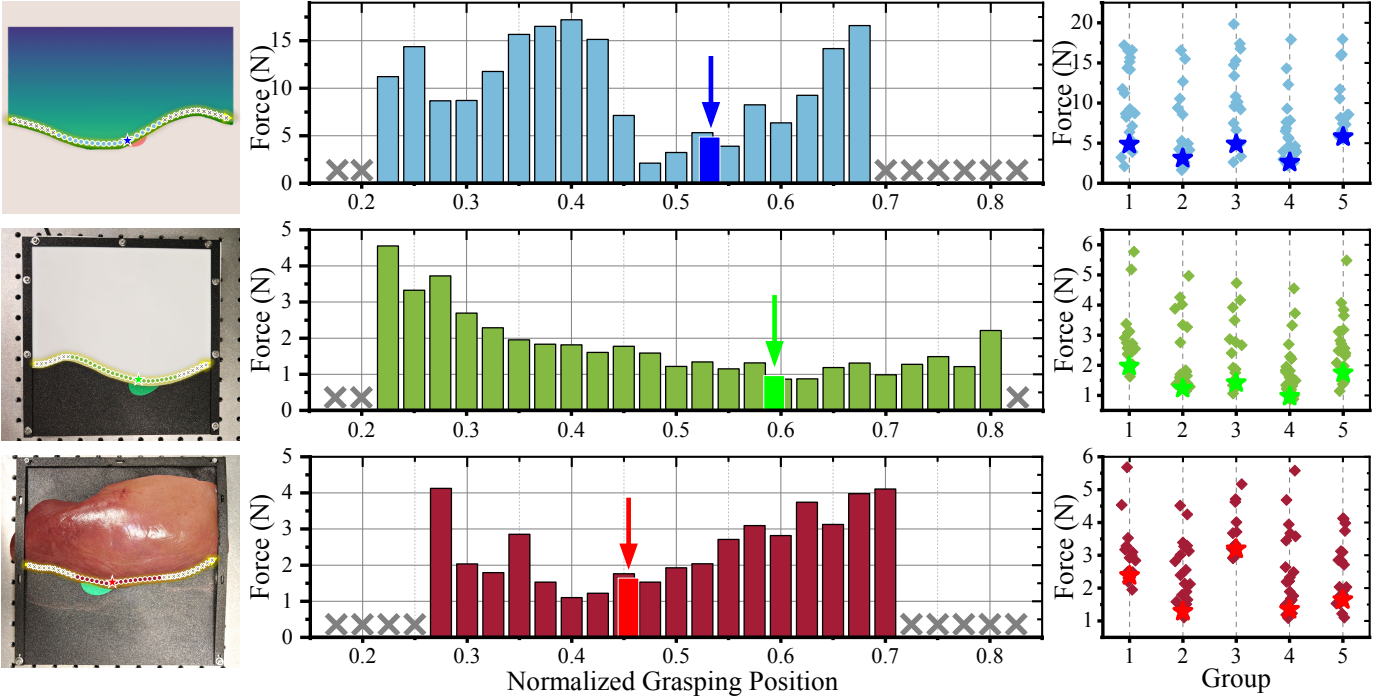


Fig. 8. (Left) Spatial distribution of candidate grasping points along the tissue boundary. (Middle) Retraction forces required to achieve full ROI exposure when executing autonomous retraction from each candidate point (lower values are better). The grasping positions are normalized from left to right. The highlighted part corresponds to the grasping point selected by the proposed algorithm, while gray crosses indicate positions where exposure could not be completed under safety constraints. (Right) Statistical summary of retraction forces across five distinct configurations for each experimental scenario.

visual boundary $E\mathcal{P}^v$ serving as a state variable in the control loop is directly extracted from the mesh nodes of the tissue model. However, in real-world experiments, it must be obtained through an additional perception pipeline. To this end, we employ Segment Anything Model 2 (SAM2) [24] to extract the visual boundary from laparoscopic images, as illustrated in Fig. 7. Specifically, SAM2 generates an image mask using a few offline-annotated point prompts, enabling the extraction of the tissue’s contour and the identification of its intersection points with the overlap area. Based on this boundary information, the optimal grasping position is computed offline via optimization and then manually set at the corresponding site prior to retraction. The controller then iteratively computes and executes motions to expose the ROI.

All experimental scenarios utilized the same deformation estimation network, trained exclusively on simulation data. However, since the closed-loop frequency and motion dynamics of the physical setup differed from those of the simulated experiments, we further adjusted the control parameters in the real-world experiments. These settings deliberately deviate from the simulation task configuration to enhance the algorithm’s performance, as detailed in Table I.

TABLE I
HYPERPARAMETER CONFIGURATION FOR EXPERIMENTS

	κ_1	κ_2	λ	s^v	s^ω	ϵ
Simulation	5e-3	1e-5	2e4	1e-3	1e-2	0.9
Real-world	1e-2	1e-5	1e5	2e-3	1e-2	0.75

Analysis of multiple retraction trials demonstrates that the proposed method completes the majority of tasks. In rare cases, such as when the ROI is placed in an extremely unfavor-

able configuration, the required deformation may exceed the maximum allowable deformation force constraint, potentially leading to exposure failure. Importantly, such failures arise from the physical infeasibility of the exposure task under those conditions, rather than from flaws in the method itself. For configurations deemed feasible a priori based on boundary geometry and ROI placement, the method achieves a success rate approaching 100%. Fig. 6 presents snapshot sequences from a representative retraction trial in each experimental scenario, showing both a global view and a laparoscopic view. The latter provides an intuitive indication of the degree of ROI exposure achieved during the procedure.

D. Grasping Point Selection

To enhance retraction performance, the proposed method incorporates an optimized selection of grasping points. This set of experiments aims to evaluate the performance of the grasping point selection algorithm. Given a specific tissue boundary shape and region-of-interest configuration, the method computes an optimal grasping position using a genetic algorithm. For comparative purposes, we uniformly sampled candidate grasping points along the tissue boundary at 5 mm intervals over the 20 cm span, and included the genetically optimized point within this candidate set.

The full autonomous retraction procedure is then executed from each candidate grasping position, and the deformation force exerted on the tissue upon complete ROI exposure is recorded. Unlike in simulation, where internal nodal forces were directly accessible, such internal forces could not be measured directly in real-world experiments. Instead, a force

sensor under the platform was used to indirectly measure the applied retraction force F^r . Since internal tissue deformation forces were generally positively correlated with the external loading, F^r was adopted in this work as a surrogate metric for tissue deformation force.

For each experimental scenario, we conducted retraction trials under five different tissue and ROI configurations, recording F^r each time the ROI was successfully exposed. Fig. 8 presents the statistical distribution of retraction forces across these five configurations and provides a detailed breakdown of F^r corresponding to all candidate grasping positions for one representative configuration.

The results show that, although the current grasping selection algorithm based on the deformation estimation network fails to precisely locate the globally optimal solution due to partial observability of the ROI and errors in network prediction results, the selected positions consistently lie within regions of relatively low retraction force. This demonstrates that, while not theoretically optimal, the proposed method effectively ensures operational feasibility and safety, making it a key enabler for stable and reliable autonomous retraction.

V. DISCUSSION AND CONCLUSION

This paper proposes a learning-based adaptive control framework for autonomous tissue retraction. The approach features a closed-loop adaptive controller that leverages the deformation Jacobian and is specifically designed for surgical exposure tasks. To improve autonomy and safety, a deformation estimation network is integrated to guide grasping point selection, determine initial configurations, and enforce safety constraints. Simulations and real-world experiments validate the effectiveness of each component, demonstrating that the proposed method achieves stable and reliable tissue retraction, showing potential for clinical applications.

As illustrated by the VPPV framework [14], learning-based strategies can generate high-level trajectories, but precise execution still relies on servo controllers to execute low-level manipulation. Task-specific control methods remain a key factor in achieving reliable embodied intelligence, serving as a bridge between high-level general-purpose cognition and low-level physical execution. Building upon the successful implementation of tissue retraction, future work will explore unified modeling and control strategies for diverse surgical manipulations. Ultimately, these task-specialized modules will be integrated with high-level task planners and cognitive reasoning systems to enable safe, interpretable, and generalizable automation of surgical procedures.

VI. ACKNOWLEDGMENT

We want to express our sincere appreciation to Wuhan United Imaging Surgical Co., Ltd. (UIS) for their generous help, not only giving us access to the MP550 robotic surgery system, but also providing constant technical support. We are particularly grateful to Dr. Xie Qiang, Dr. Yu Qiuyu, and Mr. Yang Fan for their expert guidance and hands-on assistance throughout this study.

This work was supported in part by the Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China under Grant JYB2025XDXM208, in part by the Joint Funds of the National Natural Science Foundation of China under Grant U25A600011, and in part by Hubei Science and Technology Major Program under Grants 2023BCA002 and 2024BCB009.

REFERENCES

- [1] Farshid Alambeigi, Zerui Wang, Yun-hui Liu, Russell H Taylor, and Mehran Armand. Toward semi-autonomous cryoablation of kidney tumors via model-independent deformable tissue manipulation technique. *Annals of Biomedical Engineering*, 46(10):1650–1662, 2018. doi: 10.1007/s10439-018-2074-y.
- [2] Aleks Attanasio, Bruno Scaglioni, Matteo Leonetti, Alejandro F Frangi, William Cross, Chandra Shekhar Biyani, and Pietro Valdastrì. Autonomous tissue retraction in robotic assisted minimally invasive surgery—a feasibility study. *IEEE Robotics and Automation Letters*, 5(4): 6528–6535, 2020. doi: 10.1109/LRA.2020.3013914.
- [3] Swastika Chatterjee, Soumyajit Das, Karabi Ganguly, and Dibyendu Mandal. Advancements in robotic surgery: innovations, challenges and future prospects. *Journal of Robotic Surgery*, 18(1):28, 2024. doi: 10.1007/s11701-023-01801-w.
- [4] Jiaqi Chen, Longfei Ma, Xinran Zhang, and Hongen Liao. A novel model-independent approach for autonomous retraction of soft tissue. In *12th Asian-Pacific Conference on Medical and Biological Engineering*, pages 83–90. Springer, 2024. doi: 10.1007/978-3-031-51485-2_10.
- [5] Jiaqi Chen, Guochen Ning, Longfei Ma, and Hongen Liao. Autonomous Deformable Tissue Retraction System Based on 2D Visual Representation and Asymmetric Reinforcement Learning for Robotic Surgery. *IEEE Transactions on Medical Robotics and Bionics*, 7(2):595–606, 2025. doi: 10.1109/TMRB.2025.3560399.
- [6] Aukje De Boer, Martijn S Van der Schoot, and Hester Bijl. Mesh deformation based on radial basis function interpolation. *Computers & Structures*, 85(11):784–795, 2007. doi: 10.1016/j.compstruc.2007.01.013.
- [7] Yuanming Hu, Tzu-Mao Li, Luke Anderson, Jonathan Ragan-Kelley, and Frédo Durand. Taichi: a language for high-performance computation on spatially sparse data structures. *ACM Transactions on Graphics (TOG)*, 38(6):1–16, 2019. doi: 10.1145/3355089.3356506.
- [8] Alireza Karimi, Mahdi Navidbakhsh, and Hossein Yousefi. Mechanical properties of polyvinyl alcohol sponge under different strain rates. *International Journal of Materials Research*, 105(4):404–408, 2014. doi: 10.3139/146.111036.
- [9] Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2017.
- [10] Abdelouahed Kouibia, Pedro González, Miguel Pasadas, Bassim Mustafa, Hossain Oulad Yakhlef, and Loubna

- Omri. Approximation of bivariate functions by generalized Wendland radial basis functions. *Mathematics*, 12(16):2597, 2024. doi: 10.3390/math12162597.
- [11] Romain Lagneau, Alexandre Krupa, and Maud Marchal. Active deformation through visual servoing of soft objects. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 8978–8984. IEEE, 2020. doi: 10.1109/ICRA40945.2020.9197506.
- [12] Romain Lagneau, Alexandre Krupa, and Maud Marchal. Automatic shape control of deformable wires based on model-free visual servoing. *IEEE Robotics and Automation Letters*, 5(4):5252–5259, 2020. doi: 10.1109/LRA.2020.3007114.
- [13] Jiayi Liu, Sihang Yang, Yiwei Wang, Huan Zhao, and Han Ding. Autonomous Bimanual Manipulation of Deformable Objects Using Deep Reinforcement Learning Guided Adaptive Control. In *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6904–6910. IEEE, 2025. doi: 10.1109/ICRA55743.2025.11127676.
- [14] Yonghao Long, Anran Lin, Derek Hang Chun Kwok, Lin Zhang, Zhenya Yang, Kejian Shi, Lei Song, Jiawei Fu, Hongbin Lin, Wang Wei, et al. Surgical embodied intelligence for generalized task autonomy in laparoscopic robot-assisted surgery. *Science Robotics*, 10(104): eadt3093, 2025. doi: 10.1126/scirobotics.adt3093.
- [15] Daniele Meli, Eleonora Tagliabue, Diego Dall’Alba, and Paolo Fiorini. Autonomous tissue retraction with a biomechanically informed logic based framework. In *2021 International Symposium on Medical Robotics (ISMR)*, pages 1–7. IEEE, 2021. doi: 10.1109/ISMR48346.2021.9661573.
- [16] Wouter Mollemans, Filip Schutyser, Johan Van Cleyenbreugel, and Paul Suetens. Tetrahedral mass spring model for fast soft tissue deformation. In *International Symposium on Surgery Simulation and Soft Tissue Modeling*, pages 145–154. Springer, 2003. doi: 10.1007/3-540-45015-7_14.
- [17] David Navarro-Alarcon, Yun-Hui Liu, José Guadalupe Romero, and Peng Li. Model-free visually servoed deformation control of elastic objects by robot manipulators. *IEEE Transactions on Robotics*, 29(6):1457–1468, 2013. doi: 10.1109/TRO.2013.2275651.
- [18] David Navarro-Alarcon, Hui Man Yip, Zerui Wang, Yun-Hui Liu, Fangxun Zhong, Tianxue Zhang, and Peng Li. Automatic 3-d manipulation of soft objects by robotic arms with an adaptive deformation model. *IEEE Transactions on Robotics*, 32(2):429–441, 2016. doi: 10.1109/TRO.2016.2533639.
- [19] Sachin Patil and Ron Alterovitz. Toward automated tissue retraction in robot-assisted surgery. In *2010 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2088–2094. IEEE, 2010. doi: 10.1109/ROBOT.2010.5509607.
- [20] Rosalie Plantefeve, Igor Peterlik, Nazim Haouchine, and Stéphane Cotin. Patient-specific biomechanical modeling for guidance during minimally-invasive hepatic surgery. *Annals of Biomedical Engineering*, 44(1):139–153, 2016. doi: 10.1007/s10439-015-1419-z.
- [21] Ameya Pore, Davide Corsi, Enrico Marchesini, Diego Dall’Alba, Alicia Casals, Alessandro Farinelli, and Paolo Fiorini. Safe reinforcement learning using formal verification for tissue retraction in autonomous robotic-assisted surgery. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4025–4031. IEEE, 2021. doi: 10.1109/IROS51168.2021.9636175.
- [22] Ameya Pore, Eleonora Tagliabue, Marco Piccinelli, Diego Dall’Alba, Alicia Casals, and Paolo Fiorini. Learning from demonstrations for autonomous soft-tissue retraction. In *2021 International Symposium on Medical Robotics (ISMR)*, pages 1–7. IEEE, 2021. doi: 10.1109/ISMR48346.2021.9661514.
- [23] Bing Qiao, Gang Chen, and Xiufen Ye. The research of soft tissue deformation based on mass-spring model. In *2009 International Conference on Mechatronics and Automation*, pages 4655–4660. IEEE, 2009. doi: 10.1109/ICMA.2009.5244785.
- [24] Nikhila Ravi, Valentin Gabeur, Yuan-Ting Hu, Ronghang Hu, Chaitanya Ryali, Tengyu Ma, Haitham Khedr, Roman Rädle, Chloe Rolland, Laura Gustafson, et al. Sam 2: Segment anything in images and videos. *arXiv preprint arXiv:2408.00714*, 2024.
- [25] Hamed Saeidi, Justin D Opfermann, Michael Kam, Shuwen Wei, Simon Léonard, Michael H Hsieh, Jin U Kang, and Axel Krieger. Autonomous robotic laparoscopic surgery for intestinal anastomosis. *Science Robotics*, 7(62):eabj2908, 2022. doi: 10.1126/scirobotics.abj2908.
- [26] Paul Maria Scheikl, Eleonora Tagliabue, Balázs Gyenes, Martin Wagner, Diego Dall’Alba, Paolo Fiorini, and Franziska Mathis-Ullrich. Sim-to-real transfer for visual reinforcement learning of deformable object manipulation for robot-assisted surgery. *IEEE Robotics and Automation Letters*, 8(2):560–567, 2023. doi: 10.1109/LRA.2022.3227873.
- [27] Alexander Schüßler, Christian Kunz, Rayan Younis, Benjamin Alt, Jamie Paik, Martin Wagner, and Franziska Mathis-Ullrich. Semi-Autonomous Robotic Assistance for Gallbladder Retraction in Surgery. *IEEE Robotics and Automation Letters*, 10(7):7468–7475, 2025. doi: 10.1109/LRA.2025.3577430.
- [28] Eleonora Tagliabue, Daniele Meli, Diego Dall’Alba, and Paolo Fiorini. Deliberation in autonomous robotic surgery: a framework for handling anatomical uncertainty. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 11080–11086. IEEE, 2022. doi: 10.1109/ICRA46639.2022.9811820.
- [29] Bao Thach, Brian Y Cho, Alan Kuntz, and Tucker Hermans. Learning visual shape control of novel 3D deformable objects from partial-view point clouds. In *2022 International Conference on Robotics and Au-*

- tation (ICRA), pages 8274–8281. IEEE, 2022. doi: 10.1109/ICRA46639.2022.9812215.
- [30] Bao Thach, Brian Y Cho, Shing-Hei Ho, Tucker Hermans, and Alan Kuntz. Deformernet: Learning bimanual manipulation of 3d deformable objects. *arXiv preprint arXiv:2305.04449*, 2024.
 - [31] Bao Thach, Siyeon Kim, Britton Jordan, Mohanraj Shanthi, Tanner Watts, Shing-Hei Ho, James M Ferguson, Tucker Hermans, and Alan Kuntz. DefFusionNet: Learning Multimodal Goal Shapes for Deformable Object Manipulation via a Diffusion-based Probabilistic Model. *arXiv preprint arXiv:2506.18779*, 2025.
 - [32] Mark P Wachowiak, Xiaogang Wang, Aaron Fenster, and Terry M Peters. Compact support radial basis functions for soft tissue deformation. In *2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821)*, pages 1259–1262. IEEE, 2004. doi: 10.1109/ISBI.2004.1398774.
 - [33] Holger Wendland. Error estimates for interpolation by compactly supported radial basis functions of minimal degree. *Journal of Approximation Theory*, 93(2):258–272, 1998. doi: 10.1006/jath.1997.3137.
 - [34] Bohan Yang, Bo Lu, Wei Chen, Fangxun Zhong, and Yun-Hui Liu. Model-free 3-d shape control of deformable objects using novel features based on modal analysis. *IEEE Transactions on Robotics*, 39(4):3134–3153, 2023. doi: 10.1109/TRO.2023.3269347.
 - [35] Bohan Yang, Congying Sui, Fangxun Zhong, and Yun-Hui Liu. Modal-graph 3D shape servoing of deformable objects with raw point clouds. *The International Journal of Robotics Research*, 42(14):1213–1244, 2023. doi: 10.1177/02783649231198900.
 - [36] Mingrui Yu, Kangchen Lv, Hanzhong Zhong, Shiji Song, and Xiang Li. Global model learning for large deformation control of elastic deformable linear objects: An efficient and adaptive approach. *IEEE Transactions on Robotics*, 39(1):417–436, 2023. doi: 10.1109/TRO.2022.3200546.
 - [37] Fangxun Zhong, Yaqing Wang, Zerui Wang, and Yun-Hui Liu. Dual-arm robotic needle insertion with active tissue deformation for autonomous suturing. *IEEE Robotics and Automation Letters*, 4(3):2669–2676, 2019. doi: 10.1109/LRA.2019.2913082.