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

Image-based Closed-loop Control of a Robotically Steerable Endoscopic Cannula for Minimally Invasive Neurosurgery

Nidhi Malhotra*, *Graduate Student Member, IEEE*, Revanth Konda, *Member, IEEE*, and Jaydev P. Desai, *Fellow, IEEE*

***Abstract*— Robot-assisted minimally invasive neurosurgeries have shown great promise in enabling lower invasiveness and faster patient recovery times. However, performing such surgeries remains challenging, mainly due to the use of rigid surgical tools and limited accessibility to deep-seated brain structures. Employing robotically steerable tools could address these challenges, as these devices, being relatively more dexterous, can gain access to different regions in the brain. The autonomous control of these tools could further enable manipulation with higher precision and lower procedural time, facilitating less fatigue for surgeons. In this paper, we present a control strategy for the precise manipulation of a robotically steerable endoscopic cannula (RSEC). The proposed control architecture uses a combination of inverse kinematics, endoscopic imaging, and electromagnetic tracking feedback to perform task-space control of the RSEC in real-time. A joint angle estimation algorithm is proposed to estimate the bending angles of the RSEC using an endoscopic camera. The tip-position RMSE value of the RSEC when bending the proximal and distal joints, obtained using the proposed control strategy, was 0.7 mm. The results indicate that the proposed method can be used to achieve position control of the RSEC with sub-mm accuracy.**

I. Introduction

In comparison to open neurosurgeries, employment of minimally invasive surgeries (MIS) can potentially prevent damage to eloquent brain structures due to less invasiveness [1], [2]. Treatment of intracranial hemorrhage, tumors, and hydrocephalus are common procedures wherein MIS are being adopted [3], [4]. Endoscopy is used during MIS, wherein an endoscope typically employs a trocar with a camera and a light source that may have working channel(s) for surgical cannulas to pass through. Furthermore, robotic assistance in endoscopic procedures has shown evidence of enhanced accuracy in manipulating the surgical tools [5], [6].

However, endoscopic procedures present several challenges, mainly due to the usage of rigid surgical tools, which lack the required dexterity to gain access to deep-seated brain regions. Continuum robotic tools, due to their inherent ability to generate dexterous motion, have presented a promising alternative to rigid surgical tools [7], [8]. The past few years have seen progress in the development of flexible robotically steerable endoscopic cannulas, which are a class of continuum robots applied to minimally invasive neurosurgery [9], [10]. These devices exhibit enhanced dexterity, allowing surgeons to gain access to different regions in the brain while potentially reducing tissue trauma. However, precise control of these devices is still challenging due to their flexible structure. In addition, an endoscopic camera only provides planar information and does not provide 3D information, such as depth. As a result, extensive training of the operator is often required for precise navigation in the brain. Furthermore, precise maneuverability, which results in longer surgical procedures, could lead to increased fatigue of the operator. To address these issues, research on the autonomy of sub-tasks of MIS through closed-loop control of the surgical tools has gained popularity [11], [12].

Endoscopic imaging is a popular feedback modality, frequently used in existing literature, to estimate the physical state of the device under consideration. Past studies have presented several methods to estimate the pose of rigid tools using endoscopic imaging [13], [14]. Gadwe *et al.* demonstrated real-time 6 degrees-of-freedom (DoF) pose estimation of a rigid endoscopic instrument with an outer diameter (OD) of 6.7 mm, using printable markers [15]. While this method could be utilized for pose estimation, the markers utilized in this work are difficult to incorporate into tools with a smaller OD. Furthermore, the geometric constraints to estimate the pose from the motion of the markers developed in this work do not translate to continuum robotic tools. Qin *et al.* presented instrument segmentation for a rigid tool using a fusion of the convolutional neural networks (CNNs) prediction and the kinematic pose information [16]. However, the pose estimation in real-time is not presented in this work.

Few studies have also developed methods for pose estimation of continuum tools using endoscopic imaging [17]. Cabras *et al.* proposed a color-based image segmentation with a graph-based method for instrument segmentation by detecting the color transitions along Bézier curves fitted to the tool's edges for 3D pose estimation [18]. Zhou *et al.* presented a marker-less instrument segmentation approach using network-based domain randomization [19]. Some approaches have utilized robot kinematics for pose estimation [20], [21]. Lu *et al.* showed pose estimation using a rendering approach wherein the pose is estimated by backpropagating the loss constituting the difference between the rendered image and the actual image [22]. While past studies have presented methods to obtain pose information from endoscopic imaging, limited work has been conducted on using endoscopic

The authors are with the Medical Robotics and Automation (RoboMed) Laboratory, Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA.

*Nidhi Malhotra is the corresponding author (nmalhotra37@gatech.edu).

Research reported in this publication was supported in part by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R21HD101321, and Emory University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

imaging in real-time. Consequently, closed-loop control of robotically steerable tools using endoscopic imaging has not been sufficiently explored.

In this work, we present closed-loop control of the RSEC using endoscopic imaging. The RSEC under consideration is a modified version of the robotically steerable tools developed in the prior work [9], [23]. The proposed control strategy utilizes the inverse kinematics and joint statics models of the RSEC in combination with two feedback modalities, namely 1) electromagnetic (EM) tracking, which is utilized to determine the plane of bending of the robot, and 2) endoscopic imaging, which is utilized to estimate the pose of the robot in a given plane of bending. The inverse kinematics of the RSEC are derived using optimization. The proposed control strategy is validated through a series of experiments. The primary contributions of this paper are:

- Estimation of bending angles of the RSEC using image feedback from an endoscopic camera.
- Derivation of inverse kinematics of the RSEC for task-space control.
- Closed-loop control of the joints of the RSEC using endoscopic imaging and EM tracking in real-time.

II. Materials and Methods

A. Robotically Steerable Endoscopic Cannula (RSEC) Design

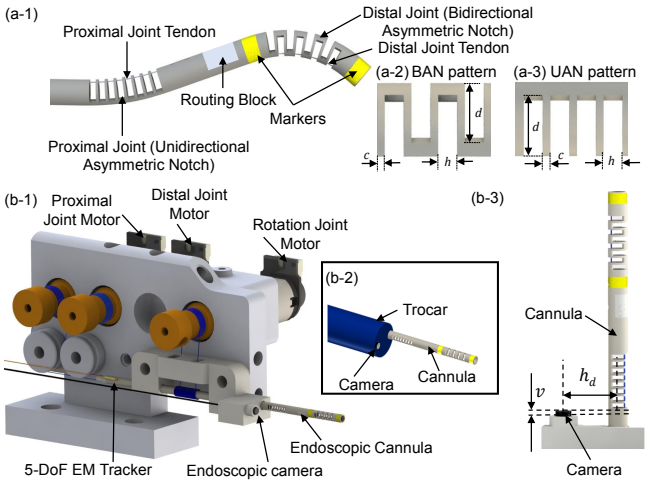

Fig. 1. Schematic of the: (a-1) robotically steerable endoscopic cannula (RSEC), (a-2) BAN pattern, and (a-3) UAN pattern. (b-1) Actuation system for the RSEC. (b-2) Schematic depicting the camera and cannula placement in a clinical scenario. (b-3) Camera offsets from the origin of the RSEC.

The design of the RSEC under consideration is shown in Fig. 1(a-1). The cannula, with an OD of 1.93 mm, consists of two steerable joints obtained by machining a nitinol (NiTi) tube using a laser micromachining platform (WS Flex Ultra-Short Pulse Laser Workstation, Optec, Frameries, Belgium). The machined segment at the distal end of the cannula, hereby referred to as the distal joint, consists of a bidirectional asymmetric notch (BAN) pattern (Fig. 1(a-2)). The machined segment at the proximal end of the cannula, hereby referred to as the proximal joint, consists of a unidirectional asymmetric notch (UAN) pattern (Fig. 1(a-3)). The notch patterns are characterized by the depth of cut (d), the notch spacing (c), the notch width (h), and the number of notches (n). The values of the notch parameters utilized to obtain the two machined segments are presented in Table I. The joints are machined to achieve bending in opposite directions with respect to each other, in the same plane. Therefore, simultaneous actuation of the joints results in S-shaped curves [23]. The "S-shape" curves can be used to avoid obstacles while being steered to the target location and can help to ensure that the distal tip of the cannula is in the field of view of the camera [9]. Furthermore, the combination of the BAN joint and a UAN joint reduces coupling between the proximal and distal joints since the actuation forces for bending the distal joint are significantly lower than the actuation forces required to cause a similar amount of bending of the proximal joint [23].

To achieve joint actuation, NiTi tendons are attached to the distal ends of each joint. Each NiTi tendon is wrapped around a spool connected to a DC motor (HPCB 6 V, Pololu Corp., NV) with a gear ratio of 380:1 for the distal joint and 1000:1 for the proximal joint. The joints are actuated by applying tension to the tendons through the actuation of motors. To achieve the rotation motion of the cannula for varying the plane of bending, an ultra-high molecular weight polyethylene (UHMWPE) tendon is wrapped around the base of the cannula and the spool of the DC motor (HPCB 6 V, Pololu Corp., NV) with a gear ratio of 380:1. The actuation system of the RSEC is shown in Fig. 1(b-1).

TABLE I
PARAMETERS OF THE RSEC.

Parameters	Value (mm)
Outer diameter, $2r_o$	1.93
Inner diameter, $2r_i$	1.49
Proximal joint, d	1.68
Proximal joint, h	0.5
Proximal joint, c	0.2
Proximal joint, n	8
Length of proximal joint, l_p	5.6
Distal joint, d	1.6
Distal joint, h	0.5
Distal joint, c	0.2
Distal joint, n	8
Length of distal joint, l_d	5.6
Length of segment between joints, l_{pd}	8.8
Length of segment after the distal joint, l_{de}	2.4
Distance between central axes of the EM-tracker and cannula, r	1.25

To control the developed RSEC in real-time, two sensing modalities are used: (a) an EM-Tracker to estimate the rotation motion, and (b) an endoscopic camera to estimate the bending angles of the RSEC. The sensor placements are shown in Fig. 1(b-1). The placement of the camera with respect to the cannula is determined based on their relative positions in a clinical setting, as shown in Fig. 1(b-2). A 5-DoF EM tracker (Aurora, Northern Digital Inc., Waterloo, Ontario, Canada) is attached to the rear end of the RSEC. The

EM tracker's position data is used to estimate the rotation angle since the angle information from the EM tracker could have only been directly utilized if the sensor was placed in line with the central line of the RSEC body. However, the EM tracker is attached to the outer surface of the RSEC body such that the central lumen of the RSEC is left open for end-effector integration, if required, which would be the case in a clinical scenario.

The endoscopic camera used in this work is the OV6946 (Omnivision, Inc., USA) [24]. The length and OD of the camera are 6 mm and 1.6 mm, respectively. The camera is characterized by 160000 pixels (400 pixels × 400 pixels), with a 120° diagonal field of view (FOV), and a 30 Hz frame rate. The camera is placed such that a region of the proximal joint and distal joint with maximum pixel area of the markers is obtained in the FOV of the camera for efficient segmentation. In this study, a simplified pin-hole camera model is assumed. The placement of the camera with respect to the cannula is characterized by the horizontal distance (h_d) and vertical distance (v) of the camera from the cannula, as shown in Fig. 1(b-3). To maintain simplicity, v was selected to be 0. The intrinsic parameters of the camera, the joint regions selected for segmentation, and the robot kinematics are utilized to determine h_d. This procedure is described in the subsection 'Joint Angle Estimation Algorithm'.

B. Joint Kinematics

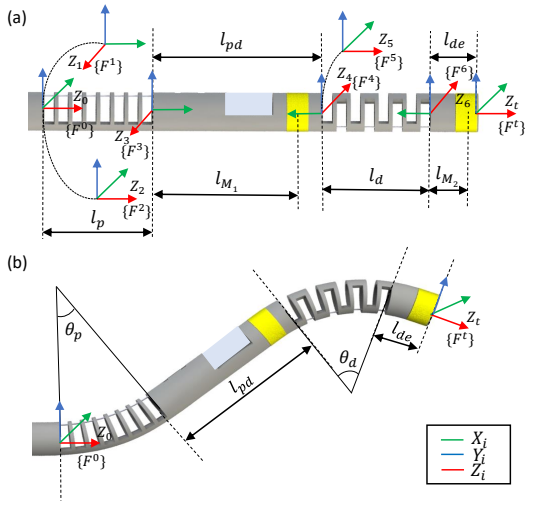

Fig. 2. Schematic of the RSEC in (a) unactuated and (b) actuated state.

1) Forward Kinematics: The motion of the RSEC can be characterized using three variables: rotation angle (θ_r), proximal bending angle (θ_p), and distal bending angle (θ_d). Each bending joint is modeled as a rotation-prismatic-rotation (RPR) joint. The origin of the world frame, $\{\mathrm{F}^0\}$, coincides with the base of the proximal joint as shown in Fig. 2. The joints (j) for the forward kinematics of the RSEC shown in Fig. 2 are defined as follows:

- j_1: Revolute joint with rotation θ_r about the Z_0-axis;
- j_2: Revolute joint with rotation $\frac{\theta_p}{2}$ about the Z_1-axis;
- j_3: Prismatic joint with translation θ_{pr_p} along the Z_2-axis;
- j_4: Revolute joint with rotation $\frac{\theta_p}{2}$ about the Z_3-axis;
- j_5: Revolute joint with rotation $\frac{\theta_d}{2}$ about the Z_4-axis;
- j_6: Prismatic joint with translation θ_{pr_d} along the Z_5-axis;
- j_7: Revolute joint with rotation $\frac{\theta_d}{2}$ about the Z_6-axis;

Assuming that the proximal and distal joints bend with constant curvatures, respectively, the prismatic motion of the proximal (θ_{pr_p}) and distal (θ_{pr_d}) joints are given by:

$$\theta_{pr_p} = \frac{2l_p}{\theta_p}\sin(\frac{\theta_p}{2}) - l_p;\ \theta_{pr_d} = \frac{2l_d}{\theta_d}\sin(\frac{\theta_d}{2}) - l_d, \quad (1)$$

where l_p and l_d denote the length of the proximal joint and distal joint, respectively. The unit angular velocity, ω, for each revolute joint and the unit linear velocity, ν, for each prismatic joint, with respect to the world frame, $\{\mathrm{F}^0\}$, is given by:

$$\omega_1 = \begin{bmatrix}0\\0\\1\end{bmatrix};\ \omega_2 = \begin{bmatrix}-1\\0\\0\end{bmatrix};\ \nu_3 = \begin{bmatrix}0\\0\\1\end{bmatrix};\ \omega_4 = \begin{bmatrix}-1\\0\\0\end{bmatrix};\ \omega_5 = \begin{bmatrix}1\\0\\0\end{bmatrix};\ \nu_6 = \begin{bmatrix}0\\0\\1\end{bmatrix};\ \omega_7 = \begin{bmatrix}1\\0\\0\end{bmatrix}, \quad (2)$$

where subscript in ω_i or ν_i corresponds to the joint j_i, with $i \in [1, 7]$. The position vector, q_i, of a point on the axis of rotation relative to the world frame, $\{\mathrm{F}^0\}$, for each revolute joint, is given by:

$$q_1, q_2 = \begin{bmatrix}0\\0\\0\end{bmatrix};\ q_4 = \begin{bmatrix}0\\0\\l_p\end{bmatrix};\ q_5 = \begin{bmatrix}0\\0\\l_p + l_{pd}\end{bmatrix};\ q_7 = \begin{bmatrix}0\\0\\l_p + l_{pd}\\ +l_d\end{bmatrix}, \quad (3)$$

where l_{pd} denotes the length of the rigid link between the proximal and distal joints. The twist coordinate, $\xi \in \mathbb{R}^6$, for a revolute joint is given by $\xi = [-\omega \times q, \omega]^{\mathrm{T}}$ and for prismatic joint is given by $\xi = [\nu, \mathbf{0}]^{\mathrm{T}}$. Using the twists, $\hat{\xi}_i \in se(3)$, the forward kinematics of the RSEC can be determined through the product of exponentials [25], given by:

$$\mathbf{g}^0_{st}(\Theta) = (\prod_{i=1}^{7} e^{\hat{\xi}_i \theta_i})\mathbf{g}^0_{st}(0) = \left[\begin{array}{c|c} \mathbf{R}^0_{st} & \begin{matrix}p_x\\p_y\\p_z\end{matrix} \\ \hline \mathbf{0}_{1\times 3} & 1 \end{array}\right] \in SE(3), \quad (4)$$

where subscript, i, in $\hat{\xi}_i$ and θ_i corresponds to j_i, $\mathbf{R}^0_{st}$ and $[p_x, p_y, p_z]^T$ is the orientation and position of the tool-frame, $\{\mathrm{F}^t\}$, given in the world frame, $\{\mathrm{F}^0\}$, respectively. The homogeneous transformation matrix relating $\{\mathrm{F}^t\}$ to $\{\mathrm{F}^0\}$, when the RSEC is unactuated, $\mathbf{g}^0_{st}(0)$, is given by:

$$\mathbf{g}^0_{st}(0) = \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & l_p + l_{pd} + l_d + l_{de}\\ 0 & 0 & 0 & 1\end{bmatrix} \in SE(3), \quad (5)$$

where l_{de} denotes the length between the distal tip of the distal joint to the tip of the cannula.

2) Inverse Kinematics: To achieve task-space control, the joint angles are identified for a desired point $P_{des}\ (p_x, p_y, p_z)$ in space with respect to the world frame, $\{\mathrm{F}^0\}$. Deriving closed-form equations for the joint angles for a given desired point is not possible for the RSEC because of the highly nonlinear nature of the forward kinematic equations.

Therefore, in this study, the inverse kinematics are derived through optimization, which is a popular technique often employed for such robots [26]. The algorithm to compute the inverse kinematics is as follows: Firstly, the plane of bending of the cannula characterized by θ_r is determined. Due to the mechanical structure of the cannula consisting of two joints bending in opposite directions, for a given (p_x,p_y), two possible values of θ_r can exist. These values are given by θ_{r_1} and θ_{r_2}, and are determined using the atan2 function.

Secondly, for θ_{r_1} and θ_{r_2}, the values of the two bending joint angles ($\alpha = [\theta_p, \theta_d]$) are determined using constrained optimization. This is accomplished by utilizing the possible values of θ_r, the desired reference point, constraints on the range of the joint angles (θ_p,θ_d), and the initial guess for the joint angles (θ_{p_0},θ_{d_0}). We used the constraint optimization solver, *fmincon* (MATLAB 2024a, The Mathworks Inc., Natick, MA) to solve for the joint angles (θ_p and θ_d) and obtain two solutions corresponding to θ_{r_1} and θ_{r_2}. The cost function, J, for the optimizer is chosen to be the square of the Euclidean distance between the desired point P_{des} and the point $\tilde{P}$ obtained for given joint values using the forward kinematics mapping. Lastly, from the two sets of solutions, the solution set that results in the minimum cost function value is selected as the desired joint angle. The algorithm is detailed in Algorithm 1 where $f(\alpha, \theta_r)$ is the forward kinematics mapping, and $\tilde{P} = [\tilde{p}_x, \tilde{p}_y, \tilde{p}_z]$, is the position for the given joint angle values. The algorithm is validated in simulation for two trajectories shown in Fig. 3, where the RMSE between the desired trajectory and the outputs of Algorithm 1 are indicated.

Algorithm 1: Inverse Kinematics with Optimization

Input: Desired P_{des} (p_x, p_y, p_z)
Output: α = [θ_p, θ_d], θ_r
$\theta_{r_1} = -\text{atan2}(p_x, p_y)$
if $p_x > 0$ **then**
$\theta_{r_2} = \pi - \text{atan2}(p_x, p_y)$
else if $p_x < 0$ **then**
$\theta_{r_2} = -(\pi + \text{atan2}(p_x, p_y))$
else $\theta_{r_2} = 0$
$J(\alpha, \theta_r) = |[\tilde{p}_x, \tilde{p}_y, \tilde{p}_z] - [p_x, p_y, p_z]|_2^2$ s.t. $\tilde{P} \leftarrow f(\alpha,\theta_r)$
$\hat{\alpha}_1 = \min_{\alpha} J(\alpha, \theta_{r_1})$ s.t. $0 < \alpha(1) < \frac{5\pi}{36}$ & $0 < \alpha(2) < \frac{\pi}{2}$
$\hat{\alpha}_2 = \min_{\alpha} J(\alpha, \theta_{r_2})$ s.t. $0 < \alpha(1) < \frac{5\pi}{36}$ & $0 < \alpha(2) < \frac{\pi}{2}$
if $J(\hat{\alpha}_1, \theta_{r_1}) \leq J(\hat{\alpha}_2, \theta_{r_2})$ **then**
| $\alpha = \hat{\alpha}_1, \theta_r = \theta_{r_1}$
else
| $\alpha = \hat{\alpha}_2, \theta_r = \theta_{r_2}$
end

C. Marker Segmentation

A marker-based approach is utilized to detect the bending angle of the flexible joints of the RSEC (see Fig. 4). The cannula was laser-micromachined with hatched regions at the distal end and the base of the distal joint. These hatched regions are filled with yellow paint for segmentation purposes. The hatched regions are segmented in the endoscopic

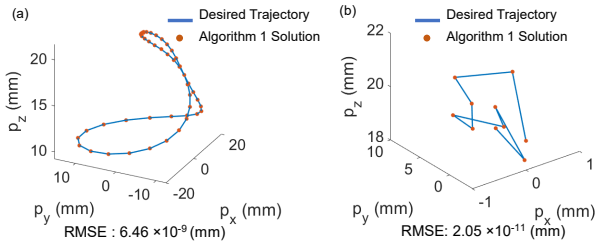

Fig. 3. Algorithm 1 validation for a desired trajectory using simulation. Results with (a) Trajectory 1, and (b) Trajectory 2.

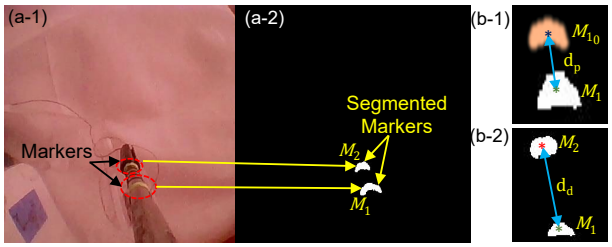

Fig. 4. (a-1) Raw image and (a-2) segmented image with proximal and distal markers of the RSEC in unactuated state. A zoomed-in view of the distance between the (b-1) proximal marker in the frame at first instance (M_{1_0}) and in the actuated state (M_1), and (b-2) proximal (M_1) and distal (M_2) markers when the distal joint is actuated.

camera image based on the Hue, Saturation, and Value (HSV) color-space segmentation thresholds. The image is corrected for distortion, binarized using the HSV threshold values, and erosion and dilation functions are used to eliminate any speckles in the image. The centroids of the two largest regions are stored as the centroids of the markers. Furthermore, only the region in the image that can contain the projection of the markers based on the workspace of the cannula is searched during the segmentation process.

D. Joint Angle Estimation Algorithm

The change in the position of the EM-tracker sensor is utilized to estimate the rotation angle, θ_r, of the cannula. When the RSEC is in its home configuration, the position value is $P_{em_i}(p_{em_{ix}}, p_{em_{iy}}, p_{em_{iz}})$. At any time instance for a position $P_{em_t}(p_{em_{tx}}, p_{em_{ty}}, p_{em_{tz}})$, the rotation angle is estimated using the following relation:

$$\theta_r = 2\text{asin}\left(\frac{||P_{em_i} - P_{em_t}||}{2r}\right)\text{sgn}(p_{em_{iz}} - p_{em_{tz}}), \quad (6)$$

where r, which is determined experimentally, is the distance between the central axes of the EM-tracker sensor and the cannula (given in Table I), and sgn denotes the signum function. It is noted that P_{em_i} and P_{em_t} are measured in the frame attached to the EM-tracker.

The camera model in combination with the forward kinematics model of the RSEC is utilized to estimate the joint angles. Utilization of the kinematic mapping of continuum robots has shown computationally efficient estimation of joint angles [27]. The marker at the base of the distal joint, 'M_1', is used to estimate the proximal joint angle. The distance between the marker at the tip of the distal joint, 'M_2', and 'M_1' is used to estimate the distal joint angle. As stated earlier, the origin of the world frame is considered

to be at the base of the proximal joint. The world frame is offset from the camera frame by a distance of h_d in the Y-direction and is rotated by an angle of θ_c about the Z-axis of the camera frame. Using the camera intrinsic and extrinsic parameters [28], the projection (u,v) of any point (p_x,p_y,p_z) on the image plane of the camera is given by:

$$\begin{bmatrix} u*k \\ v*k \\ 1*k \end{bmatrix} = \begin{bmatrix} f_x & S & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c(\theta_c) & -s(\theta_c) & 0 & 0 \\ s(\theta_c) & c(\theta_c) & 0 & h_d \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_y \\ p_z \\ 1 \end{bmatrix}, \tag{7}$$

where k denotes the scale, f_x and f_y denote the focal length of the camera along its X-axis and Y-axis respectively, S denotes the skew, (u_0, v_0) denote the pixel coordinates of the optical center of the camera, $c(\cdot)$ and $s(\cdot)$ are the cosine and sine functions, respectively. Since, only a rotation about the Z-axis is assumed, $k = p_z$. The coordinates of the centroid of the markers in the image are denoted as (u_{m_1},v_{m_1}) for M_1 and (u_{m_2},v_{m_2}) for M_2.

1) Proximal Joint Angle Estimation: To estimate θ_p using endoscopic imaging, firstly, a 2D map is created for the change in position of M_1 with respect to its initial position. The initial position of M_1 is obtained when the robot is in its unactuated state. It is noted that the surface of the cannula that is closer to the camera is visible in the image plane. The camera is offset in the negative Y-direction with respect to the world frame, $\{\mathrm{F}^0\}$. Therefore, the portion of the marker visible in the image plane is offset in the negative Y-direction from the world frame by the outer radius of the cannula. This offset is considered while using the forward kinematics to compute the position of the markers. The 2D map, computed for different values of θ_p and θ_r, utilizes the forward kinematic equations to determine the position as follows: Based on the respective ranges, θ_p and θ_r are discretized using a 1° step-size. The range of θ_r is $[-180^\circ, 180^\circ]$ and θ_p is $[0^\circ, 25^\circ]$. For each value of θ_r and θ_p, the position of the centroid of M_1, denoted by $P_{M_1} = [P_{M_1,x}, P_{M_1,y}, P_{M_1,z}]^T$, is computed using the forward kinematics and is given by the following equation:

$$\begin{bmatrix} P_{M_1,x} \\ P_{M_1,y} \\ P_{M_1,z} \end{bmatrix} = \begin{bmatrix} s(\theta_r)(l_{M_1} s(\theta_p) + l_p \frac{1-c(\theta_p)}{\theta_p} - r_o c(\theta_p)) \\ c(\theta_r)(l_{M_1} s(\theta_p) + l_p \frac{1-c(\theta_p)}{\theta_p} - r_o c(\theta_p)) \\ l_{M_1} c(\theta_p) + l_p \frac{s(\theta_p)}{\theta_p} + r_o s(\theta_p) \end{bmatrix}, \tag{8}$$

where r_o and l_{M_1} denote the outer radius of the cannula and distance from the center of the M_1 to the distal end of the proximal joint, respectively (shown in Fig. 2).

Using Eq. (7), the different position coordinates of M_1 are projected onto the camera frame. For each projected coordinate, the Euclidean distance between the projected coordinates of the robot in its unactuated state, and its current state is computed and is denoted by $\hat{d}_p$. This process is accomplished before the start of the experiment. The generated 2D distance projection map is stored as a matrix and is used as a look-up table to determine the measured value of θ_p.

Secondly, the experimental value of θ_r is determined based on the EM-tracker position data. We assume that the joint is in its unactuated state at time t_0, when the first frame is captured, and the centroids for M_1 are given by ($u_{M_{1_0}}$,$v_{M_{1_0}}$). Thirdly, at any given instant of time, the projection of the centroids for M_1, given by (u_{M_1},v_{M_1}) are determined. This is followed by computing the Euclidean distance between the initial position of the centroid and the current position of the centroid, d_p, as shown in Fig.4 (b-1) using the following equation:

$$d_p = \sqrt{(u_{M_1} - u_{M_{1_0}})^2 + (v_{M_1} - v_{M_{1_0}})^2}. \tag{9}$$

Finally, for a given value of d_p and θ_r, the value of θ_p that results in $\hat{d}_p \approx d_p$, is selected from the 2D map, as the measured proximal joint angle.

2) Distal Joint Angle Estimation: Similar to the proximal joint estimation, to estimate θ_d, firstly, a look-up table is created. However, unlike the proximal joint estimation, which utilizes a 2D map with $\hat{d}_p$ values, the distal joint estimation uses a 3D map, which is generated using the discretized values of θ_r, θ_p, and θ_d. For each of these discretized values, the position of the centroid of M_1 is computed as stated above. The position of the centroid of M_2, denoted by P_{M_2}, is computed using forward kinematics and is given as follows:

$$\begin{bmatrix} P_{M_2,x} \\ P_{M_2,y} \\ P_{M_2,z} \end{bmatrix} = \begin{bmatrix} s(\theta_r)(l_p \frac{1-c(\theta_p)}{\theta_p} + l_{pd} s(\theta_p) + l_d(\frac{c(\theta_d-\theta_p)-c(\theta_p)}{\theta_d}) \\ -l_{M_2} s(\theta_d - \theta_p) - r_o c(\theta_d - \theta_p)) \\ c(\theta_r)(l_p \frac{1-c(\theta_p)}{\theta_p} + l_{pd} s(\theta_p) + l_d(\frac{c(\theta_d-\theta_p)-c(\theta_p)}{\theta_d}) \\ -l_{M_2} s(\theta_d - \theta_p) - r_o c(\theta_d - \theta_p)) \\ l_p \frac{s(\theta_p)}{\theta_p} + l_{pd} c(\theta_p) + l_d(\frac{s(\theta_p)+s(\theta_d-\theta_p)}{\theta_d}) \\ +l_{M_2} c(\theta_d - \theta_p) + r_o s(\theta_d - \theta_p) \end{bmatrix}, \tag{10}$$

where l_{M_2} denotes the distance from the distal end of the distal joint to the center of M_2 as shown in Fig. 2.

These coordinates are projected onto the camera frame using Eq. (7). The Euclidean distance between the projected coordinates of M_1 and M_2, denoted by $\hat{d}_d$, are computed and used to generate the 3D distance projection map. This process is accomplished before the start of the experiment. The generated 3D map is stored as a matrix and is used as a look-up table to determine the measured value of θ_d. Secondly, the experimental values of θ_r and θ_p are determined based on the EM-tracker position data and endoscopic imaging, respectively. Thirdly, at any given instant of time, the projection of the centroids of M_1 and M_2, given by (u_{M_1},v_{M_1}) and (u_{M_2},v_{M_2}), respectively, are determined. This is followed by computing the Euclidean distance between the position of the centroids of M_1 and M_2, d_d, as shown in Fig. 4 (b-2) using the following equation:

$$d_d = \sqrt{(u_{M_2} - u_{M_1})^2 + (v_{M_2} - v_{M_1})^2}. \tag{11}$$

Finally, for a given value of d_d, θ_r, and θ_p, the value of θ_d that results in $\hat{d}_d \approx d_d$, is selected from the 3D map, as the measured distal joint angle.

The horizontal distance h_d, shown in Fig. 1(b-2), is chosen such that the markers are visible in the imaging plane. When the camera is placed away from the cannula, the segmented region of the markers increases in the image plane. However,

since the distance between the camera and the cannula increases, the markers move away from the optical center of the image, thereby going out of the FOV of the camera. For a given h_d, M_1 and M_2 lie farthest from the camera, when $\theta_p = \theta_{p_m}(\theta_{p_m} = \max(\theta_p))$, $\theta_r = 0$ and $\theta_c = 0°$. To obtain a simplified expression to estimate h_d, we neglect the skew in the intrinsic camera matrix. The projection of the centroid of M_1 and M_2 can be given by $(0,v_{M_1})$ and $(0,v_{M_2})$, respectively, and v_{M_1} and v_{M_2} can be given by:

$$\frac{|v_{M_1} - v_0|}{f_y} = \frac{|(P_{M_1,y}) + h_d|}{P_{M_1,z}}; \frac{|v_{M_2} - v_0|}{f_y} = \frac{|(P_{M_2,y}) + h_d|}{P_{M_2,z}}, \tag{12}$$

where h_d is chosen to be the maximum value for which v_{M_1} and v_{M_2} lie within the image plane.

E. Control Strategy

Firstly, the RSEC is controlled in an open-loop manner to reach a desired point, P_{des}, with respect to the world frame of the robot, $\{F^0\}$. For this purpose, the inverse kinematics mapping is utilized to obtain the desired joint angles: θ_r, θ_p, and θ_d for a given P_{des} in the workspace of the robot. The joint statics models for the rotation joint, UAN joint, and the BAN joint derived in [9] are used for obtaining the desired tendon stroke values for the motor inputs. The joint statics models only consider the loading cycle of the joints and neglect the friction and hysteresis. Therefore, feedback is used to further enhance the control of the RSEC (see Fig. 5).

Secondly, the RSEC is controlled in a closed-loop manner by utilizing feedback to correct for the joint angles. The rotation angle, θ_{r_m}, is estimated using Eq. (6) using the EM-tracker feedback, and the bending joint angles (α_m) are estimated using the joint angle estimation algorithm. In addition to the inverse model, a proportional gain controller is used to correct the error in the estimated and desired joint angle values obtained from the inverse kinematics mapping. The proportional gain of the controller is tuned heuristically.

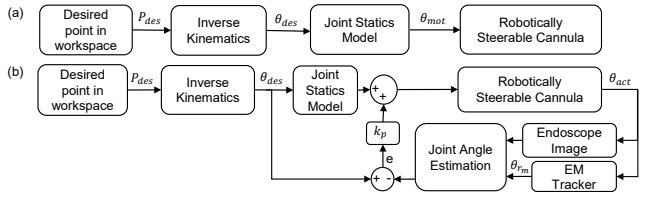

Fig. 5. Block diagram for (a) open-loop, and (b) closed-loop control. P_{des}, θ_{des}, θ_{mot}, θ_{act}, α_m, and θ_{r_m}, are the desired position in the workspace of the RSEC, the desired joint angles, the motor inputs, the actual joint values, the bending joint angles estimated by the joint angle estimation algorithm, and the measured rotation angle from the EM-tracker, respectively.

III. Results

A. Camera Calibration

The calibration for the endoscopic camera was performed using a checkerboard pattern. A total of 105 images were grabbed, and Camera Calibration Toolbox (MATLAB 2024a) was used to obtain the camera parameters. The intrinsic matrix for the camera was determined to be as follows:

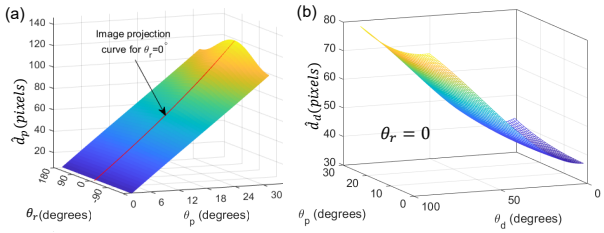

Fig. 6. Distance map for estimating (a) θ_p and (b) θ_d (at $\theta_r = 0$).

$$K = \begin{bmatrix} 260.4 & -0.2965 & 191.65 \\ 0 & 259.9 & 199.57 \\ 0 & 0 & 1 \end{bmatrix} \tag{13}$$

The maximum value of h_d using the intrinsic parameters, θ_{p_m} = 25°, and v_{M_1} = 5 pixels (chosen > 0 pixel such that a considerable part of the marker is within the image), is computed to be 5.63 mm and for v_{M_2}= 5 pixels is computed to be 7.55 mm from Eq. (12). To ensure that both the markers appear in the image with at least 5 pixels, the smaller value of the two obtained values of h_d is selected. To account for any unmodeled phenomenon, h_d was chosen to be 5.5 mm.

B. Validation of Joint Angle Estimation Algorithm

To validate the joint angle estimation algorithm, firstly, we generate the 2D map for the proximal joint and the 3D map for the distal joint based on the possible joint values as shown in Fig. 6. Secondly, we evaluate the algorithm for different trajectories. To do so, the bending angles are measured using a 6-DoF EM-tracker (Aurora, Northern Digital Inc., Waterloo, Ontario, Canada) attached to the tip of the cannula, which are considered to be the ground truth values (θ_r(EM), θ_p(EM), θ_d(EM)). It is noted that the 6-DoF EM-tracker is used only for the purpose of validating the proposed control strategy and is not intended for inclusion in the robotic system. The joint angles (θ_p(IM), θ_d(IM)) are estimated using the image feedback and the 5-DoF EM-tracker system. The proximal joint estimation is verified for $\theta_r \approx 0$ in Fig. 7(a) and a varying rotation angle in Fig. 7(b). The distal joint estimation is verified for $\theta_r \approx 0$ and $\theta_p \approx 0$ in Fig. 7(c), for $\theta_r \approx 0$ and a varying θ_p in Fig. 7(d) and for $\theta_p \approx 0$ and a varying θ_r in Fig. 7(e). The root-mean-square-error (RMSE) values for joint angle estimation for the five trajectories are given in Table II.

C. Control Experiments

To validate the proposed control strategy, we evaluate the RSEC to track the desired trajectories in the cannula's workspace. To estimate the desired joint angles (θ_{des}) required for reaching a desired point P_{des} in space, we utilize the inverse kinematics mapping. Furthermore, the desired joint angles are used to compute the desired tendon stroke values using the joint statics models. In addition, to correct for the errors due to the unmodeled effects in the statics models, the EM-tracking and imaging feedback are utilized with the developed models to achieve closed-loop control.

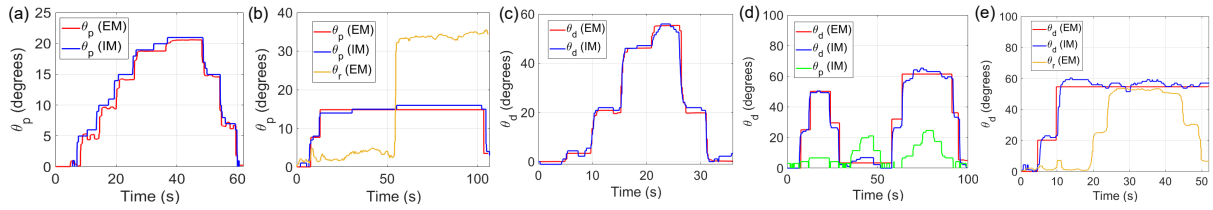

Fig. 7. Validation experiments for Joint Angle Estimation Algorithm. Validation of proximal joint estimation for: (a) $\theta_r \approx 0$, and (b) varying θ_r. Validation of distal joint estimation for: (c) $\theta_r \approx 0$ and $\theta_p \approx 0$, (d) $\theta_r \approx 0$ with varying θ_p, and (e) $\theta_p \approx 0$ with varying θ_r.

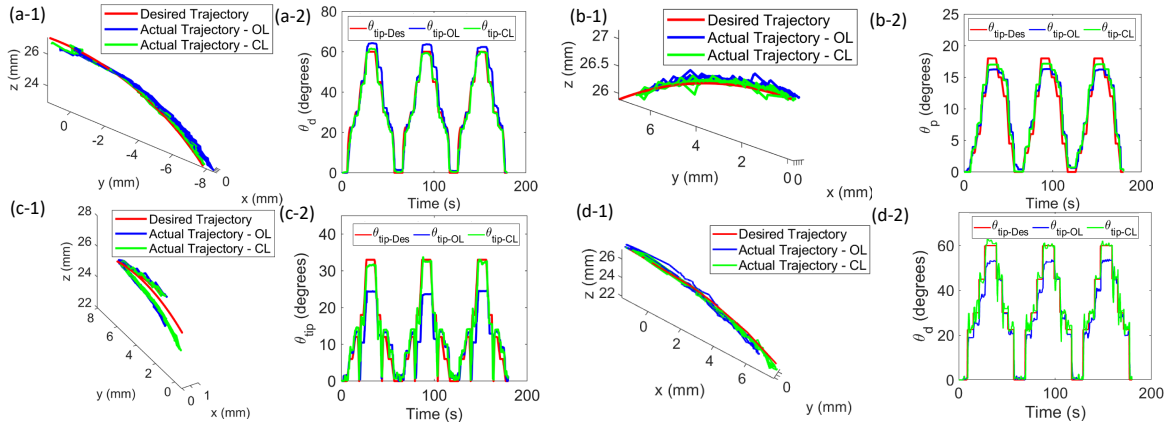

Fig. 8. Control experiments for: (a-1) Trajectory 1: The desired trajectory, actual trajectory for open-loop case (Actual Trajectory-OL), and closed-loop case (Actual Trajectory-CL) are shown. (a-2) Trajectory 1: The desired bending joint angle at the tip ($\theta_{\text{tip-Des}}$), the actual joint angle for open-loop ($\theta_{\text{tip-OL}}$) and closed-loop ($\theta_{\text{tip-CL}}$) control are shown. Similar plots are shown for: (b-1)-(b-2) Trajectory 2, (c-1)-(c-2) Trajectory 3, and (d-1)-(d-2) Trajectory 4.

The performance of the controller in open-loop and closed-loop is shown in Fig. 8, and the RMSE value for the error in the position measured by the EM-tracker and the desired position is given in Table III. The desired ($\theta_{\text{tip-Des}}$) and the measured bending joint angles at the tip of the cannula are shown in Fig. 8 for each trajectory for the open-loop ($\theta_{\text{tip-OL}}$) and the closed-loop ($\theta_{\text{tip-CL}}$) experiments. It is to be noted that the tip positions shown in Fig. 8 are the positions of the center of the EM tracker placed at the tip used for validation; thus, the desired positions in the trajectories are calculated by adding the length from the cannula tip to the center of the tracker to the overall length of the cannula (l_{de} is adjusted to be 7.4 mm). The tip bending angle of the cannula θ_{tip} is given by $|\theta_p - \theta_d|$. The RMSE value between the desired tip angle and the measured tip angle for the given trajectory is given in Table III. The experiments show sub-mm accuracy in reaching the desired positions using the proposed strategy. Prior literature showed tip position errors of 4.18 mm, 1.24 mm, and 1.29 mm, in [18], [19], [21], respectively with real-time performance shown in [19] and [21] wherein the computation times for pose estimation were stated as 800 ms and 41 ms.

D. Discussion

Using the proposed control strategy, the experiments exhibited sub-mm accuracy in reaching the desired positions in the workspace of the cannula. Furthermore, the search for the estimated bending joint angles using the joint angle estimation algorithm is achieved within approximately 3 ms. The experiments were performed using a computer with 32 GB RAM, which runs at 3 GHz speed. The processing time is within the frame rate of the camera used in this work, thereby facilitating real-time control. Additionally, for rotation angle estimation, another sensor such as a hollow shaft encoder, can be utilized at the base of the cannula to measure the rotation. However, irrespective of the sensing method used for the estimation of θ_r, the image-based control algorithm should remain the same.

TABLE II

RMSE FOR VALIDATION OF JOINT ANGLE ESTIMATION

Fig. No.	Fig. 7(a)	Fig. 7(b)	Fig. 7(c)	Fig. 7(d)	Fig. 7(e)
RMSE	1.12°	1.63°	2.74°	4.60°	3.08°

TABLE III

CONTROL EXPERIMENTS RMSE-VALUES.

Fig. No.	Fig. 8(a)	Fig. 8(b)	Fig. 8(c)	Fig. 8(d)
P_{des} RMSE-OL	0.82 mm	0.53 mm	1.06 mm	0.93 mm
P_{des} RMSE-CL	0.49 mm	0.35 mm	0.70 mm	0.73 mm
$\theta_{\text{tip-Des}}$ RMSE-OL	4.3°	1.7°	5.4°	4.3°
$\theta_{\text{tip-Des}}$ RMSE-CL	3.1°	1.4°	2.4°	2.5°

A few aspects of the system that can be further improved are as follows. Firstly, the segmentation approach utilized in the control strategy may not perform efficiently in the presence of occlusion, drastic changes in the lighting, or other similar phenomena that may appear in clinical settings. This can be addressed by using machine learning techniques for robust segmentation or by combining the proposed approach with other sensing methods [29]. Secondly, the camera utilized in this work has a limited FOV that restricts

the workspace in which the cannula can be controlled. This is similar to clinical scenarios, where the cannula is often controlled by the clinician only in the FOV of the camera. This issue can be circumvented by utilizing a camera with enhanced FOV. Lastly, a proportional gain controller was employed in this work, which exhibited acceptable performance. However, the control performance can be further enhanced by utilizing advanced control strategies such as adaptive control [30] or sliding mode control [31].

IV. Conclusion

This paper presents a control strategy for a robotically steerable endoscopic cannula. The proposed strategy utilizes imaging feedback from an endoscopic camera and EM-tracking feedback. An algorithm to estimate the joint angles using endoscopic feedback is proposed. Furthermore, an inverse kinematic model of the RSEC, derived using optimization, is used to compute the desired joint angles for the controller. The proposed control method exhibited acceptable performance in tracking the desired tip position of the RSEC. The methods presented in this work can potentially be used for other robotically steerable tools for real-time control. Future work will focus on adapting the proposed method to a neuroendoscopic tool consisting of two cannulas and utilizing sensor fusion approaches employing shape estimation through strain sensing, along with the proposed approach.

References

[1] S. Shaikh and C. Deopujari, "The endoscope and instruments for minimally invasive neurosurgery," *Mini-invasive Surgery*, vol. 4, p. 89, 2020.

[2] M. Ishii and G. L. Gallia, "Application of technology for minimally invasive neurosurgery," *Neurosurgery Clinics*, vol. 21, no. 4, pp. 585–594, 2010.

[3] A. V. Kulkarni, J. M. Drake, C. L. Mallucci, S. Sgouros, J. Roth, S. Constantini, C. P. N. S. Group *et al.*, "Endoscopic third ventriculostomy in the treatment of childhood hydrocephalus," *The Journal of pediatrics*, vol. 155, no. 2, pp. 254–259, 2009.

[4] J. P. Greenfield and M. M. Souweidane, "Endoscopic management of intracranial cysts," *Neurosurgical focus*, vol. 19, no. 6, pp. 1–9, 2005.

[5] M. Iftikhar, M. Saqib, M. Zareen, and H. Mumtaz, "Artificial intelligence: revolutionizing robotic surgery," *Annals of Medicine and Surgery*, vol. 86, no. 9, pp. 5401–5409, 2024.

[6] P. Łajczak, K. Żerdziński, K. Jóźwik, M. Laskowski, and M. Dymek, "Enhancing Precision and Safety in Spinal Surgery: A Comprehensive Review of Robotic Assistance Technologies," *World Neurosurgery*, 2024.

[7] J. Rosen, L. N. Sekhar, D. Glozman, M. Miyasaka, J. Dosher, B. Dellon, K. S. Moe, A. Kim, L. J. Kim, T. Lendvay, Y. Li, and B. Hannaford, "Roboscope: A flexible and bendable surgical robot for single portal Minimally Invasive Surgery," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, 2017, pp. 2364–2370.

[8] P. Theiss and A. Alaraj, "Commentary: Preclinical Evaluation of a Novel Steerable Robotic Neuroendoscope Tool," *Operative Neurosurgery*, vol. 26, no. 4, pp. 396–397, 2024.

[9] R. Qi, N. Malhotra, T. A. Brumfiel, K. Hoang, and J. P. Desai, "Development of a single port dual arm robotically steerable endoscope for neurosurgical applications," *npj Robotics*, vol. 3, no. 1, p. 1, 2025.

[10] K. W. Eastwood, T. Looi, H. E. Naguib, and J. M. Drake, "Design optimization of neuroendoscopic continuum instruments for third ventriculostomy and tumor biopsy," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 4853–4856.

[11] A. Attanasio, B. Scaglioni, E. De Momi, P. Fiorini, and P. Valdastri, "Autonomy in surgical robotics," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, no. 1, pp. 651–679, 2021.

[12] H. Saeidi, H. N. D. Le, J. D. Opfermann, S. Leonard, A. Kim, M. H. Hsieh, J. U. Kang, and A. Krieger, "Autonomous Laparoscopic Robotic Suturing with a Novel Actuated Suturing Tool and 3D Endoscope," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 1541–1547.

[13] M. K. Hasan, L. Calvet, N. Rabbani, and A. Bartoli, "Detection, segmentation, and 3D pose estimation of surgical tools using convolutional neural networks and algebraic geometry," *Medical Image Analysis*, vol. 70, p. 101994, 2021.

[14] X. Du, T. Kurmann, P.-L. Chang, M. Allan, S. Ourselin, R. Sznitman, J. D. Kelly, and D. Stoyanov, "Articulated Multi-Instrument 2-D Pose Estimation Using Fully Convolutional Networks," *IEEE Transactions on Medical Imaging*, vol. 37, no. 5, pp. 1276–1287, 2018.

[15] A. Gadwe and H. Ren, "Real-Time 6DOF Pose Estimation of Endoscopic Instruments Using Printable Markers," *IEEE Sensors Journal*, vol. 19, no. 6, pp. 2338–2346, 2019.

[16] F. Qin, Y. Li, Y.-H. Su, D. Xu, and B. Hannaford, "Surgical instrument segmentation for endoscopic vision with data fusion of cnn prediction and kinematic pose," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 9821–9827.

[17] D. Bouget, M. Allan, D. Stoyanov, and P. Jannin, "Vision-based and marker-less surgical tool detection and tracking: a review of the literature," *Medical Image Analysis*, vol. 35, p. 633–654, 2017. [Online]. Available: https://api.semanticscholar.org/CorpusID:206870493

[18] P. Cabras, F. Nageotte, P. Zanne, and C. Doignon, "An adaptive and fully automatic method for estimating the 3D position of bendable instruments using endoscopic images," *The International Journal of Medical Robotics and Computer Assisted Surgery*, vol. 13, no. 4, p. e1812, 2017.

[19] C. Zhou, L. Wang, B. Wu, and K. Xu, "A Markerless 3D Tracking Framework for Continuum Surgical Tools Using a Surgical Tool Partial Pose Estimation Network Based on Domain Randomization," *Advanced Intelligent Systems*, vol. 6, no. 4, p. 2300434, 2024.

[20] R. Reilink, S. Stramigioli, and S. Misra, "3D position estimation of flexible instruments: marker-less and marker-based methods," *International journal of computer assisted radiology and surgery*, vol. 8, pp. 407–417, 2013.

[21] R. A. Porto, F. Nageotte, P. Zanne, and M. d. Mathelin, "Position control of medical cable-driven flexible instruments by combining machine learning and kinematic analysis," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 7913–7919.

[22] J. Lu, F. Liu, C. Girerd, and M. C. Yip, "Image-based pose estimation and shape reconstruction for robot manipulators and soft, continuum robots via differentiable rendering," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 560–567.

[23] Y. Chitalia, S. Jeong, J. Bok, V. Nguyen, S. Melkote, J. J. Chern, and J. P. Desai, "Towards the Design and Development of a Pediatric Neuroendoscope Tool," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 2998–3004.

[24] OmniVision, "OV6946," [Accessed: Oct. 1, 2024]. [Online]. Available: https://www.ovt.com/products/ov6946/

[25] R. M. Murray, Z. Li, and S. S. Sastry, *A mathematical introduction to robotic manipulation*. CRC press, 2017.

[26] R. Qi, N. U. Nayar, and J. P. Desai, "Compact Design and Task Space Control of a Robotic Transcatheter Delivery System for Mitral Valve Implant," *IEEE Transactions on Medical Robotics and Bionics*, 2023.

[27] N. U. Nayar and J. P. Desai, "Ultrasound-Guided Real-Time Joint Space Control of a Robotic Transcatheter Delivery System," *IEEE Robotics and Automation Letters*, vol. 9, no. 9, pp. 7677–7684, 2024.

[28] R. Szeliski, *Computer vision: algorithms and applications*. Springer Nature, 2022.

[29] A. K. Rothe, N. Malhotra, C. C. Barré, and J. P. Desai, "Towards shape estimation of meso-scale continuum robots using direct laser written piezoresistive strain sensors," in *2025 International Symposium on Medical Robotics (ISMR)*, 2025, pp. 108–114.

[30] I. D. Landau, R. Lozano, M. M'Saad, and A. Karimi, *Adaptive control: algorithms, analysis and applications*. Springer Science & Business Media, 2011.

[31] J.-J. E. Slotine, W. Li *et al.*, *Applied nonlinear control*. Prentice hall Englewood Cliffs, NJ, 1991, vol. 199, no. 1.