IEEE Robotics and Automation Letters (RAL) paper presented at
2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

EyeLS: Shadow-Guided Instrument Landing System for Target Approaching in Robotic Eye Surgery

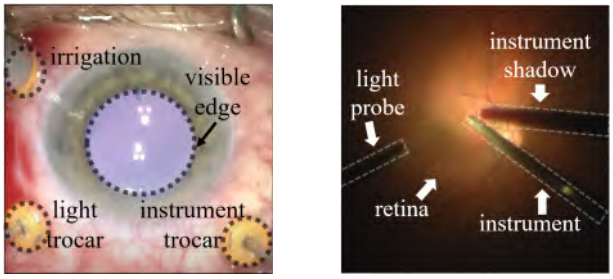
Junjie Yang[1], Zhihao Zhao[1], Siyuan Shen[1], Daniel Zapp[2], Mathias Maier[2], Kai Huang[3], Nassir Navab[1], *Fellow, IEEE*, and M. Ali Nasseri[2*]

***Abstract*—Robotic ophthalmic surgery is an emerging technology to facilitate high-precision interventions such as subretinal injection and removing swinging tissues in retinal detachment using microscopy and iOCT. However, locating the instrument tip outside iOCT's range-limited ROI is challenging, especially at the initial target-approaching stage. Meanwhile, due to 2D perspective projection and the lack of depth perception with the required micron precision, current image-based methods cannot effectively navigate the instrument tip's trajectory towards both intra-retinal and above-retinal target points. To address this limitation, we propose using shadows of the instrument tip and target to estimate their relative depth position and optimize the instrument tip's insertion trajectory until it approaches targets within the iOCT's scanning area. Our method achieves a mean depth error of 0.0127 mm for above-retinal targets and 0.3473 mm for intra-retinal targets in the surgical simulator without damaging the retina, triggering subsequent iOCT-dominant micron-precision manipulations. This method also succeeds in the experiment of target approaching on a retina model.**

***Index Terms*—Medical Robots and Systems, Vision-Based Navigation, Visual Servoing**

I. Introduction

In ophthalmic surgeries, surgeons pursue micron-scale precision for intraocular status estimation and avoiding anatomic damage during instrument-tissue interaction. As a promising high-precision solution, intraoperative optical coherence tomography (iOCT) is integrated with microscopy in the Operation Room (OP) [1]–[3], especially for the subretinal injection procedure [4]. However, the iOCT-guided surgical applications, such as instrument-pose estimation and retina reconstructions [5]–[11], take effect only when the surgical instrument moves within the iOCT's limited scanning range (commonly 6x6 mm^2 square area). Although novel methods, such as spectrally encoded reflectometry (SER) [12] and 4D OCT, are invented to assist the dynamic iOCT-instrument coordination, they still need time-consuming data processing and device validation to achieve a mature surgical integration. Therefore, it is necessary to investigate more efficient image-guided methods to navigate instruments at the initial target-approaching stage without updating the current surgical hardware in the OP.



(a) Example of trocars [13]. (b) Example of surgery frames.

Fig. 1: Example of surgical setup in real surgeries.

In microscope-guided surgeries, as shown in Fig. 1, the core surgical tasks include extracting the relative instrument-target position to optimize the instrument trajectory. Few researchers propose to use laser and structured light to cast a visible light marker on the retina for the instrument navigation [14]–[19]. However, apart from their potential light toxicity, such light-based marker generation can only specify targets on the retina and, therefore, is limited in posterior-segment interventions. Other methods use microscopy to estimate the absolute position of instruments inside the eye, such as stereo-camera-based depth estimation [20] with the difficulty of calibrating the stereo camera system, and deep learning [21]–[23] to implicitly generate motion commands with the challenge of collecting diverse microscope image datasets for network training.

An inspiring light-probe control method [24] is proposed to move the light probe so that the instrument and its shadow are separated in the image during the retina-approaching procedure, hence allowing qualitative estimation of the instrument-retina distance. However, this method only considers target points on the retina and has the drawback of the instrument's long-time visual blockage of the retinal target. Therefore, it causes difficulty in monitoring the instrument-tissue interaction and affects the subsequent instrument control. Furthermore, the instrument's approaching direction towards the retinal target is limited due to its vertical tip trajectory, which

Manuscript received: 08, 11, 2023; Revised 12, 07, 2023; Accepted 02, 16, 2024.

This paper was recommended for publication by Editor Jessica Burgner-Kahrs upon evaluation of the Associate Editor and Reviewers' comments.

This work was supported by Bayerische Forschungsstiftung (AZ-1503-21).

[1]Junjie Yang, Zhihao Zhao, Siyuan Shen and Nassir Navab (Faculty) are with School of Computation, Information and Technology, Technical University of Munich, 80333 Munich, Germany {junjie.yang, zhihao.zhao, siyuan.shen, nassir.navab}@tum.de

[2]Mathias Maier, Daniel Zapp and M. Ali Nasseri are with the Klinik und Poliklinik für Augenheilkunde, Klinikum rechts der Isar, 81675 Munich, Germany {mathias.maier, daniel.zapp}@mri.tum.de, ali.nasseri@tum.de

[3]Kai Huang (Faculty) is with School of Computer Science and Engineering, Sun Yat-Sen University, 510006 Guangzhou, China huangk36@mail.sysu.edu.cn

Digital Object Identifier (DOI): see top of this page.

is not consistent with the common axial-oriented instrument insertion.

In this paper, relying on the basic principles of light and shadow, we propose and prove a theorem that given a properly-placed light probe, a desired planar point and its shadow in the microscope image determine a unique 3D point inside the eyeball. Based on this theorem, the instrument tip's 3D trajectory that intersects the target is represented by the projections of the predicted tip trajectory, passing the planar target point, and its shadow trajectory, passing the planar target shadow. Specifically, horizontal alignment and vertical-angle optimization contribute to the cannula tip's landing onto the target from the side direction by overlapping the tip and target and their shadows, respectively.

The proposed method has the listed advantages over other state-of-the-art methods: 1) A unified theory to handle both intra-retinal and above-retinal intraocular target points during relative-depth estimation; 2) Explainable and intuitive utilization of shadows for motion control; 3) The side-approaching strategy avoids visual blockage and is consistent with the direction of instrument insertion.

II. Explanation of Concept

This paper considers the surgical instrument a cannula since it covers most surgical cases. Notations for context reference are listed in TABLE. I.

TABLE I: Notations and thresholds.

variable	meaning
$\mathbb{P}$	a point set
$\mathcal{P}$	a plane
l_o / p_o	fitted line / point of object o in the image
$_{lp}$ / $_c$ / $_{cs}$ / $_t$ / $_{ts}$	light / cannula / its shadow / target / its shadow
p_{crcm}	cannula's RCM point (trocar)
$l_{p_1 \to p_2}$	line defined by two points p_1 and p_2
l^v	line that perpendicular to imaging plane
$\mathcal{J}_c = \Delta_{R/V/H}$	periodical step motion along spherical axes
$d_{o_1 \to o_2}$	distance between object o_1 and o_2
$\vec{v}_{o_1 \to o_2}$	vector from o_1 to o_2
σ_{close}	pixel threshold for enabling shadows
$\sigma_{align}^{ang/dis}$	angle / pixel-threshold of horizontal alignment
σ_{app}	pixel threshold for checking object overlapping

A. Priors and Assumptions

In ophthalmic surgeries, especially posterior-segment interventions, the patient is under general anesthesia as a prior, hence allowing the surgeon to maintain the eyeball's stability and visual alignment using eyeball-orbital control [25]. Furthermore, according to the instrument motion constrained by trocars, the cannula's trocar p_{crcm} in the image is approximated as the intersection of multiple cannula lines $[l_c^1, ..., l_c^i]$ with different horizontal angles, which is solved as a least-square problem. Image segmentation is a prior technique in this paper that helps capture and assess the dynamic intraocular environment.

In this paper, our method relies on the following assumptions: 1) The whole surgical cannula is modeled as a cylinder; 2) The retinal surface is not severely deformed to enable the cannula tip's shadow approximation; 3) The microscope and the cannula's polar plane are both vertically aligned with the eye. 4) The light probe is appropriately placed to cast brightness and shadow inside the visible range of the retina.

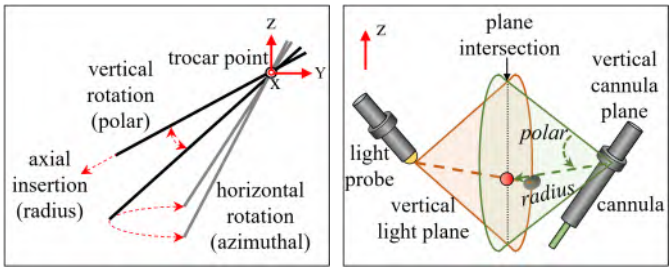
B. Target Definition

The above-retinal target points $p_t \in \mathbb{P}_{floating}$ are partially-detached retina tissues floating or swinging in the hollow vitreous space to be peeled or sucked, hence owning the feature that target p_t and its shadow p_{ts} on the retina are separated in the microscope image. In realistic surgeries, we define the desired cannula-tissue collision point as the target point p_t (usually an edge point of the tissue). The target point's shadow p_{ts} is approximated as the intersection of the light-target line $l_{lp \to t}$ and the tissue's shadow edge on the retina.

The intra-retinal target points $p_t \in \mathbb{P}_{retinal}$ are the desired retinal location on or beneath the retina for surgical manipulations, i.e., subretinal injection, hence owning the feature that p_t and p_{ts} are approximately overlapped in the microscope image even considering the retinal transparency.

C. RCM Modeling

Since trocars are placed through the conjunctiva and sclera to provide tunnels of penetration as depicted in Fig. 1(a), the control of instruments should be pivoted around trocars, which forms the Remote Center of Motion (RCM) constraint. Then, the cannula's motion is described by its radius, horizontal (azimuthal) and vertical angles (polar) in a spherical coordinate system with its origin point at the trocar, as shown in Fig. 2(a). Consequently, the target-approaching procedure is described as finding the cannula's ideal insertion trajectory that intersects the target as shown in Fig. 2(b).



(a) Spherical motion modeling. (b) Vertical target exploration.

Fig. 2: Spherical modeling for RCM and target exploration.

III. Method

A. Method Overview

The task flow of our proposed target-approaching method is depicted in Fig. 3. The horizontal alignment (b) and enabling shadow (c) execute once and are maintained during runtime. The axial insertion (d) and shadow alignment (e) are coordinated in each frame-processing loop until the cannula tip reaches the target point.

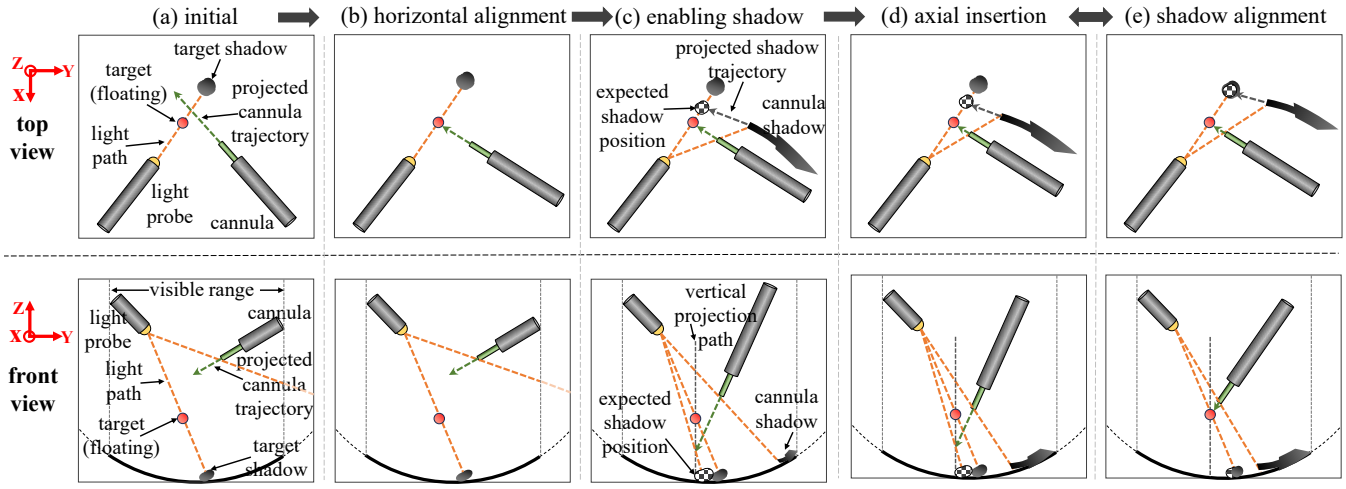


Fig. 3: Task flow of the proposed method with the comparison of the top and front view.

B. Explanation of Steps

1) Horizontal Alignment: The initial horizontal alignment is to rotate the cannula horizontally so that the cannula's projected trajectory intersects the target, at least in the image, which does not guarantee the intersection in the 3D space. Considering the cannula's segmented direction fluctuation due to unstable image segmentation, the horizontal alignment is divided into two stages based on the cannula-target distance $d_{c\to t}$ compared with threshold σ_{close}. When $d_{c\to t} > \sigma_{close}$, the intersection angle between $\vec{v}_{crcm\to c}$ and $\vec{v}_{crcm\to t}$ is maintained within σ^{ang}_{align} to finish a rough angle-based horizontal alignment due to the distance redundancy of safe cannula-target interaction. If $d_{c\to t} \leqslant \sigma_{close}$, the distance $d_{p_t\to l_c}$ is maintained within σ^{dis}_{align} to achieve precise horizontal alignment. In each frame, rotation by step angle $\pm\Delta_H$ is conducted, its direction depending on the relative cannula-target position.

2) Enabling Shadow: Suppose no cannula shadow is visible in the microscope. In this case, the cannula should be rotated toward the retina (Δ^{down}_V) within its vertical exploration plane, as shown in Fig. 2(b). Although the cannula shadow appears around the cannula tip, the consequent visual fallback of the cannula's projection should be compensated for by its axial insertion to ensure the cannula tip and its shadow are always inside the visible retinal range.

3) Axial Insertion: Since the cannula is already heading to the target in the image with both their shadows visible on the retina, the cannula tip should get closer to the target point by axial movement to achieve the target approaching. Meanwhile, the initial cannula placement and the relative cannula-target position determine the direction of axial movement. If $d_{crcm\to c} > d_{crcm\to t}$, the cannula is over-inserted, hence in need of axial fallback Δ^{out}_R to get closer to the target. Otherwise, the cannula can be further axially inserted Δ^{in}_R toward the target by increasing its insertion radius if $d_{crcm\to c} < d_{crcm\to t}$.

4) Shadow Alignment: After the cannula is horizontally aligned toward the target, as shown in Fig. 2(b), the remaining task is to find a proper polar angle for the cannula in its vertical exploration. This vertical alignment is solved by searching for the ideal tip position and trajectory.

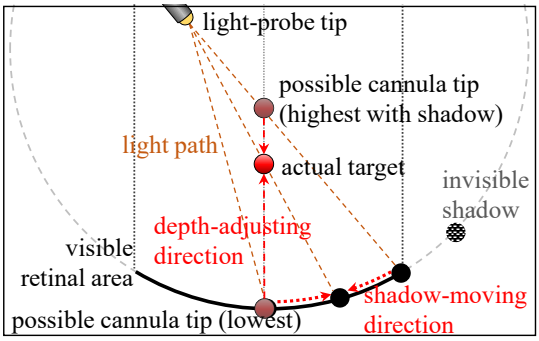


Fig. 4: Depth adjustment and shadow movement.

As for exploring the ideal tip position, the cannula tip p_c and the target point p_t are already visually overlapped. In this case, p_{lp}, p_c, p_{cs}, p_t and p_{ts} are all inside the vertical light plane $\mathcal{P}_{vlp}$ (the orange sector area in Fig. 2(b) and also in Fig. 4) that is perpendicular to the microscope's imaging plane $\mathcal{P}_{img}$.

Then, the vertical distribution of the cannula tip (with visible shadows) on the target's vertical projection line $l^v_t \in \mathcal{P}_{vlp}$ is denoted as a sequence of m cannula-tip points $\mathbb{P}_{cv} = [p^0_c, ..., p^i_c], i \in [0, m)$ in the order of increasing cannula-tip depth. Using the retina bottom as the reference position, p^0_c is the highest possible cannula-tip position floating inside the vitreous space with a visible shadow, while p^{m-1}_c is the deepest of possible cannula-tip position with its invisible shadow overlapped with the target on the retina. Subsequently, with $\mathcal{E}$ the spherical surface of the eyeball, the visible shadow point p^i_{cs} of any cannula tip $p^i_c \in \mathbb{P}_{cv}$ can be obtained using Equation (1) as a one-to-one mapping.

$$\forall p^i_c \in \mathbb{P}_{cv},\ p^i_{cs} = l_{lp\to p^i_c} \cap \mathcal{E},\ p^i_{cs} \in \mathcal{P}_{vlp} \tag{1}$$

Since all possible cannula shadow tips p_{cs} are distributed within the vertical light plane $\mathcal{P}_{vlp}$, the cannula tip's depth distribution is consistent with its shadow distribution as Equation (2) where dp means the object's depth. Given two different points $p^i_c,\ p^j_c \in \mathbb{P}_{vlp}$,

$$dp^i_c > dp^j_c \mapsto \begin{cases} d_{p_t\to p^i_{cs}} < d_{p_t\to p^j_{cs}} \text{ if } p_{lp} \text{ invisible} \\ d_{p_{lp}\to p^i_{cs}} < d_{p_{lp}\to p^j_{cs}} \text{ if } p_{lp} \text{ visible} \end{cases} \tag{2}$$

Based on Equation (2), when the cannula tip is overlapped with the target point in the image, the relative position between p_{ts} and p_{cs} can be referred to for the cannula tip's depth adjustment by achieving the overlapping of the cannula tip and the target and their shadows, as Equation (3).

$$p_c^{3d} \approx p_t^{3d} \text{ if } p_c \approx p_t \text{ and } p_{cs} \approx p_{ts} \tag{3}$$

However, the pure vertical alignment strategy, as presented in [24], has the drawback of visually blocking the target. Therefore, exploring the ideal tip position is extended into exploring the ideal tip trajectory. Since the cannula is already horizontally aligned toward the target, the cannula tip's axial insertion trajectory $\mathcal{T}_c$ is determined to intersect with the target's vertical projection line l_t^v at the expected cannula-tip position p_{ecp}^{3d}. Although p_{ecp} in the image cannot be visually recognized due to the perspective projection, its corresponding shadow point can represent it as the intersection of the line $l_{lp\to t}$ and the cannula's shadow-tip trajectory $\mathcal{T}_{cs}$ based on the shadowing principle. We define this shadow point as the expected shadow position p_{esp}, which is mapped to p_{ecp} according to Equation (1). Then, the cannula-target alignment in the 3D space as Equation (3) can be approximated as Equation (4).

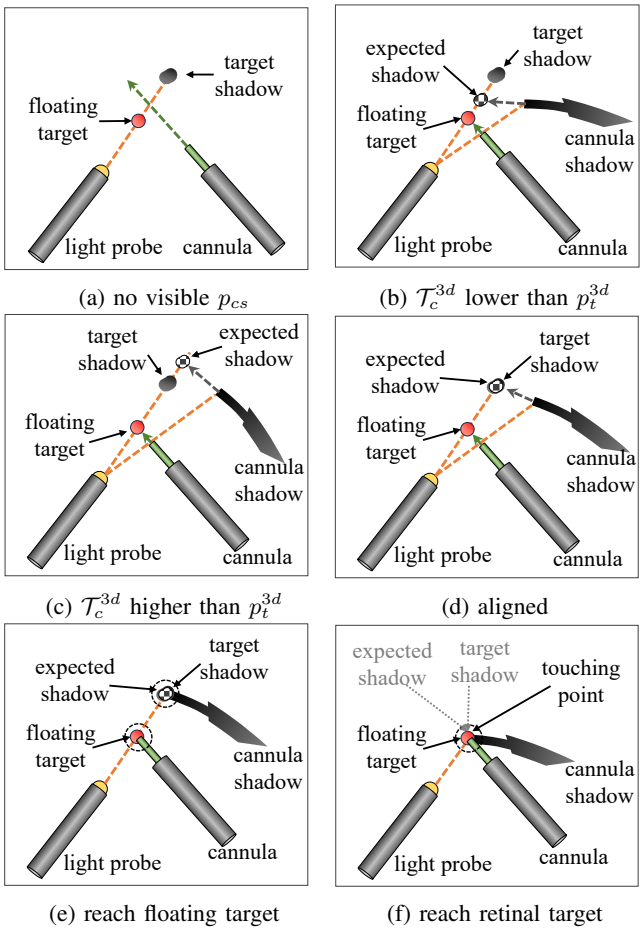
$$p_c^{3d} \approx p_t^{3d} \text{ if } p_{ecp} = p_c \approx p_t \text{ and } p_{esp} \approx p_{ts} \tag{4}$$

In practice, the cannula's future shadow-tip trajectory is predicted by extending the current shadow along its current orientation as $\mathcal{T}_{cs} \approx \mathcal{T}_{cs}^{pred} \approx l_{cs}$. Then, combining Equation (2) and (4), the simultaneous shadow alignment is achieved by considering the cases in Fig. 5(b)-(f). The corresponding motion-command generation is presented in **Algo 1**.

The horizontal alignment and its real-time maintenance are bare assertions as **Algo-1** line 1. After the recognition of point p_{lp}, p_{ts} and p_{esp}, the distance from p_{lp} to the latter two points can be calculated as $d_{lp\to ts}$ and $d_{lp\to esp}$.

Algorithm 1 Procedure of Shadow Alignment

1: AssertOrientationAligned()
2: $p_{esp} = l_{cs} \cap l_{lp\to t}$
3: **if** $d_{ts\to esp} \leqslant \sigma_{app}$
4: **if** $(d_{c\to t} \leqslant \sigma_{app})$ **and** $(d_{cs\to ts} \leqslant \sigma_{app})$
5: StartiOCT()
6: **else**
7: **if** $d_{c\to cs} \leqslant \sigma_{app}$
8: $\mathcal{J}_c = \Delta_R^{out}$ & Δ_V^{up}
9: **else**
10: **if** $d_{crcm\to c} < d_{crcm\to t}$
11: $\mathcal{J}_c = \Delta_R^{in}$
12: **else**
13: $\mathcal{J}_c = \Delta_R^{out}$
14: **else**
15: **if** $d_{lp\to ts} < d_{lp\to esp}$
16: $\mathcal{J}_c = \Delta_V^{down}$
17: **else**
18: $\mathcal{J}_c = \Delta_V^{up}$
19: MoveCannula($\mathcal{J}_c$, Δ_R)



(a) no visible p_{cs} (b) $\mathcal{T}_c^{3d}$ lower than p_t^{3d} (c) $\mathcal{T}_c^{3d}$ higher than p_t^{3d} (d) aligned (e) reach floating target (f) reach retinal target

Fig. 5: Exclusive top views for trajectory optimization.

If all $d_{ts\to esp}$, $d_{c\to t}$ approximately reach the threshold σ_{app}, both cannula tip and its shadow are correctly aligned around their target positions, hence triggering the subsequent high-precision iOCT navigation, as **Algo-1** lines 3-5. Otherwise, the shadow-aligned cannula should conduct its axial insertion/fallback by step length Δ_R depending on if the tip is beyond the target or falling behind referring to p_{crcm}, as lines 10-13 and in Fig. 5(d).

If the estimated shadow position p_{esp} is not overlapped with the target shadow p_{ts}, the relationship between $d_{lp\to ts}$ and $d_{lp\to esp}$ defines the relative cannula-target position. If $d_{lp\to ts} < d_{lp\to esp}$ as shown in Fig. 5(c), the cannula tip's insertion trajectory (referring to p_{ecp}) is above the target's 3D position, hence requiring the cannula's downwards vertical rotation downwards as **Algo-1** lines 15-16. By contrast, if $d_{lp\to ts} > d_{lp\to esp}$ as Fig. 5(b), the cannula-tip's insertion trajectory is below the target's 3D position, leading to necessary upwards vertical rotation as **Algo-1** lines 17-18.

It is worth noticing that the cannula approaches the retinal surface too fast in some situations before reaching the target point. In order to eliminate the risk of damaging non-ROI retinal areas, a safety-check step with the highest priority of execution is needed as lines 7-8 to move the cannula tip away from the retina by upwards vertical rotation and axial fallback, which ensures p_c and p_{cs} does not overlap.

IV. Experiment And Results

The experiments are divided into two categories: simulation and model test. For simulation, the automatic motion test is conducted in a Unity-based simulator to prove the correctness of our decision-making. A retina model is 3D-printed for the model test to run the target-approaching algorithm with the necessary segmentation metrics provided.

A. Simulation

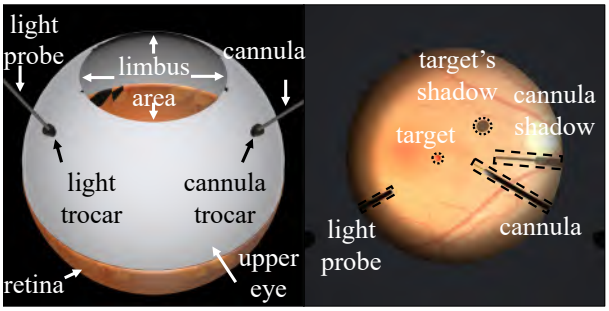


Fig. 6: The simulator used for our experiment.

1) Setup: As shown in Fig. 6, a Unity-based simulator is developed by our lab and used to experiment with cannula navigation towards intraocular targets. During the simulator development, authentic surgical images are used to synthesize simulator scenarios, and therefore, realistic surgical scenes similar to typical surgeries are generated. The light probe is statically positioned to ensure adequate brightness and correct shadow shaping. The radius of the simulated spherical eye is 12 mm with a 6mm-radius visible range.

In this experiment, the cannula-target distance is presented as the performance metric of our target-approaching algorithm. Meanwhile, the cannula-retina distance is the safety metric to demonstrate the ability to avoid tissue damage.

Here are the hyperparameters used in the simulation: the threshold for angle-based alignment σ_{align}^{ang}=3°, the threshold for distance-based alignment σ_{align}^{dis}=2 pixels, the threshold to start distance-based alignment σ_{close}=100 pixels, the threshold to trigger iOCT navigation σ_{app}=15 pixels, the step angle for rotation $\Delta_{V/H}$=0.5° and the step length for zooming Δ_R=0.067 mm. The motion steps are tuned in some cases to avoid stuck movement. Since the simulator can provide the cannula's projected tip and orientation vector, the cannula's size specification is ignored during the simulation.

TABLE II: Information of generated targets in simulation.

type	percent	* dis-x	* dis-y	* dis-z
floating	46.78%	0.60 (3.00)	-1.00 (4.71)	-9.02 (1.98)
retinal	53.22%	-0.02 (5.03)	0.06 (7.08)	-11.48 (0.08)
* Axis distribution (mm) consists of mean (variance).				

2) Result: During the simulation, we generate **1926** random intraocular targets with their mean (variance) along each axis in TABLE II, all targets within the visible range of the retina. Simulated images are shaped as 1024x1024 pixels.

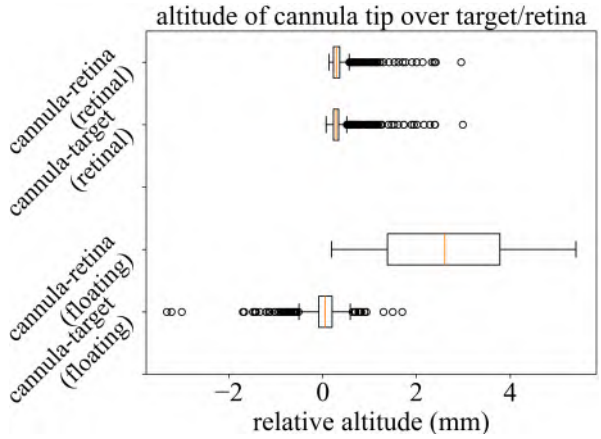


Fig. 7: The altitude of cannula tip referring to target/retina to demonstrate the performance/safety of target approaching.

With only one failure, as shown in Fig. 9, and another three stuck cases with tunable hyperparameters, the final relative altitude of the cannula tip referring to targets and retina after approaching is collected together with the vertical distance from the cannula to the retinal surface in Fig. 7. Their mean are values presented in Table III.

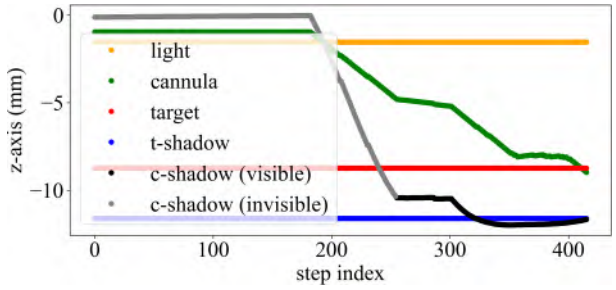
TABLE III: Cannula tip's relative altitude (mm).

target type	floating		retinal	
reference	target	retina	target	retina
mean	0.0127	2.6161	0.3473	0.3628

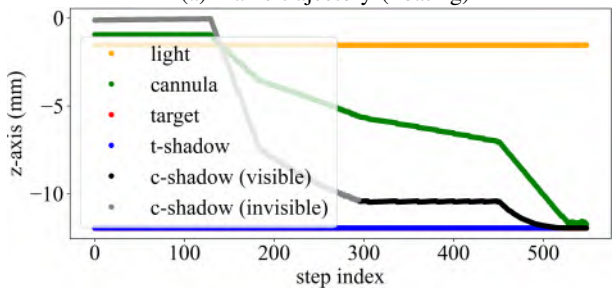
As depicted in Fig. 7, when floating targets are focused on, although the cannula-tip position is distributed around the target with both higher and lower situations with its average depth error around 0 mm, the cannula-retina relative altitude is always larger than 0 mm and within 4 mm, marking that the cannula tip is always above the retinal surface without unexpected touching. Similarly, as for retinal targets, the cannula never touches the retinal surface when approaching retinal targets. Also, since retinal targets are located on the retinal surface, the cannula's relative altitude distributions to the target and the retina are consistent. Therefore, this relative altitude distribution shows the safe cannula placement of our method without damaging the retinal tissue. At the same time, according to the detailed relative altitude values in Table. III, the mean value of cannula-target distance along the z-axis shows that the final cannula-target distance is smaller than 0.5 mm, which proves the feasibility of our image-based target-approaching method in ophthalmic surgeries. Also, the mean value of cannula-retina distance along the z-axis larger than 0.1 mm proves our algorithm's safety by avoiding unexpected cannula-retina collision, which prevents unnecessary tissue damage. Thanks to the inspiration of using cannula-shadow tip collision to predict retina touching in [24], we achieve damage-free cannula-tip control.

Furthermore, Since the typical scanning diameter of iOCT is 6 mm, the final approaching errors within 0.8 mm along both the x-axis and y-axis for both retinal and floating targets prove that the approaching method in this paper can adequately

place the cannula tip inside the iOCT's scanning range for later micron-level manipulation. Combining this error distribution in the x-y plane with the relative altitude distribution along the z-axis (depth) in Fig. 7 and Table. III, the 3D error distribution proves that our proposed target-approaching method enables the cannula tip to approach floating and retinal targets within 1 mm and allows the subsequent cannula manipulations such as insertion and sucking without damaging intraocular tissues.



(a) z-axis trajectory (floating)



(b) z-axis trajectory (retinal)

Fig. 8: The simulated cannula-tip's trajectory along z axis (depth) towards the floating and retinal targets, respectively.

Fig. 8 shows the trajectories of approaching two types of targets along the z-axis (depth) to help understand how the cannula tip is getting close to the target. As for the floating target in Fig. 8(a), the cannula first adjusts its orientation towards the target by horizontal rotation between step-index $[0, 200]$, which does not change the cannula tip's depth according to the spherical modeling, and then optimize the cannula-tip's depth to approach the target between step-index $[200, 300]$. Simultaneously, coordinated axial insertion and vertical rotation (upwards) help maintain the cannula's depth but adjust the shadow's position until both the cannula tip and its shadow tip overlap with the target's.

As for the retinal target in Fig. 8(b), the z-axis (depth) trajectory also follows the predefined step division with the target overlapped with its shadow on the retina of a 12mm-radius eye. Similar horizontal alignment and depth optimization are also achieved in the retinal-target case. However, according to the depth trajectory, a unique depth fluctuation of the cannula's shadow is observed between step-index $[300, 460]$, which is caused by the emergent retinal-collision avoidance to ensure that the cannula tip will not damage the retina before surgical tasks. Afterward, the cannula makes a smooth axial insertion towards the target to finally approach it. Since floating targets are consistently above the retinal surface and separated from their shadows, such collision avoidance does not exist in the floating-target scenario.

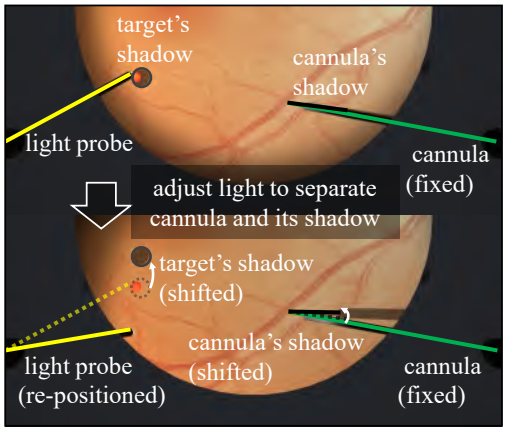


Fig. 9: The problematic case with additional light adjustment.

The case in Fig. 9 shows the navigation failure due to the visual stacking of the cannula and its shadow. Since the improper light-tip placement makes p_{lp}, p_c and p_{cs} almost on the same line and hence causes the invisible p_{ns} and failed calculation of p_{esp} for trajectory prediction. The method in [24] shows a potential solution to this challenge: manually rotating the light probe to separate the cannula and its shadow. After the target shadow and the cannula shadow are both visible, the trajectory optimization procedure continues.

B. Model Test

1) Setup: Target-approaching experiments are also conducted on a retina model. Due to the lack of depth-accessible sensors, this model-based experiment only provides qualitative effectiveness proof of our shadow-alignment algorithm. The hardware setup is shown in Fig. 10, which contains a light probe (left), retina model (bottom), microscope (up), and cannula mounted on a hybrid-structure robot [26] (right).

This paper uses a Geuder cannula (model G-34285, 23 gauge, typically used in retina treatment) for image segmentation and the subsequent target approaching. An RCM-compatible actuator listens to the image-based controller for motion commands and actuates specific motors. For the simplicity of implementation, a loose robot initialization is achieved to roughly place the robot under the microscope without strict hand-eye calibration, which is compensated by the offset-based control strategy. The cannula's shadow has already been activated in the visible area. We define the following runtime hyper-parameters: the threshold for angle-based alignment σ_{align}^{ang}=10°, the threshold for distance-based alignment σ_{align}^{dis}=30 pixels, the threshold to start distance-based alignment σ_{close}=100 pixels, the threshold to check object overlapping σ_{app}=45 pixels, step angle for rotation $\Delta_{V/H}$=0.05° and step length for insertion Δ_R=0.01 mm.

As for the recognition of image components, the light probe and the target are static objects that can be manually marked in the image. Also, two types of virtual targets are abstracted and manually set in the microscope image. The virtual target and its virtual shadow are overlapped at pixel $[512, 512]$ for the retinal case and separated as pixels $[512, 512]$

and $[671, 391]$ for the floating case in the image. Meanwhile, since no ground-truth positions of dynamic objects are given during the experiment, the cannula and its shadow positions are segmented in each frame as the decision-making input. Therefore, we train an **YOLOv8-lseg** model [27] with 158 images, 126 images for training and 32 images for validation. The segmentation model's final metrics after training are available in TABLE. IV running on a hardware combination of a 6-core CPU and an RTX-A5000 GPU. After segmentation to

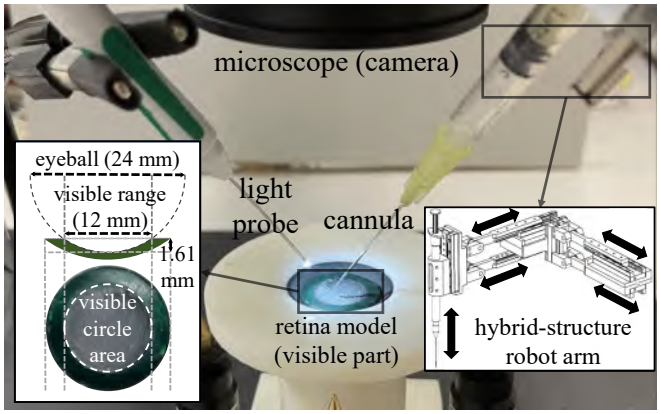


Fig. 10: The setup for experiments using a robot [26].

obtain the contours of instruments, we model the instrument's orientation by the fitted center line of its contour and use the intersection of the center line and the contour box as the tip point. This paper considers the segmented result as prior input data using existing industrial segmentation tools.

2) Result: As a qualitative test, the procedures for approaching two types of targets are presented in Fig. 11, (a)-(d) for the retinal target and (e)-(h) for the floating target. Here, the target (black) and its shadow (black) are overlapped in the retinal case and are separated in the floating case. After horizontal alignment, the cannula is controlled with axial insertion to move the shadow tip towards the target shadow along the ideal shadow trajectory until the target and cannula tip overlap, their shadow points as well. With the predefined approaching threshold σ_{app}=45 pixels, the cannula-target error, as well as their shadow error, is smaller than 0.9 mm for the retinal case and 0.837 mm for the floating case in the image plane (x-y), respectively.

TABLE IV: Training metrics.

Precision	Recall	mAP50	mAP50-95
0.981	0.985	0.984	0.837
543 epochs in total with batch size 8.			

During the model test, a sporadic shifting of the cannula tip within a small area is observed due to unstable mask generation during image segmentation. Therefore, segmentation tools should be continuously updated to prevent instability or failed segmentation during the proposed shadow utilization.

V. Discussion and Conclusion

This paper tries to loosen the constraint of vertical-only cannula movement in [24], and unifies the shadow utilization for the cannula-trajectory optimization for both floating and retinal intraocular targets. This shadow-based navigation concept is tested on a simple 3D-printed retina model, which only considers the retina's spherical surface. However, several factors in real eye surgeries influence the performance of our shadow-based method, including light-probe positioning, retina status and microscope imaging, necessitating additional experiments in a more realistic intraocular environment.

This paper only considers the light probe's illumination task to ensure the shadow's visibility. The proposed method utilizes the shadow orientation to predict and optimize its trajectory, which extends the limited usage of shadow tip for collision avoidance in [24]. Shadow alignment also allows the target approaching with more practical cannula movements than vertical-only movement. However, the shadow alignment requires proper casting of both target and cannula shadows in the visual range. It also relies on the shadow shaft's recognizability and segmentability, except for its tip visibility along the cannula movement. Such shadow-quality control is challenging, especially considering complicated intraocular optical attributes for shadow casting. Therefore, the light probe's deterministic role of shaping the shadow brings the need to develop advanced light-controlling algorithms. Furthermore, the proposed method cannot handle large-range subretinal hemorrhage and retina deformation that prevents the extraction of shadow information. Lastly, higher-resolution CMOS sensors can be mounted on the microscope lens to increase the pixel granularity for capturing detail features. Consequently, this paper encourages continuously updating SOTA segmentation tools to improve the corresponding segmentation precision.

References

[1] S. Dehghani, M. Sommersperger, P. Zhang, A. Martin-Gomez, B. Busam, P. Gehlbach, N. Navab, M. A. Nasseri, and I. Iordachita, "Robotic navigation autonomy for subretinal injection via intelligent real-time virtual ioct volume slicing," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 4724–4731.

[2] M. Zhou, X. Guo, M. Grimm, E. Lochner, Z. Jiang, A. Eslami, J. Ye, N. Navab, A. Knoll, and M. A. Nasseri, "Needle detection and localisation for robot-assisted subretinal injection using deep learning," *CAAI Transactions on Intelligence Technology*, 2023.

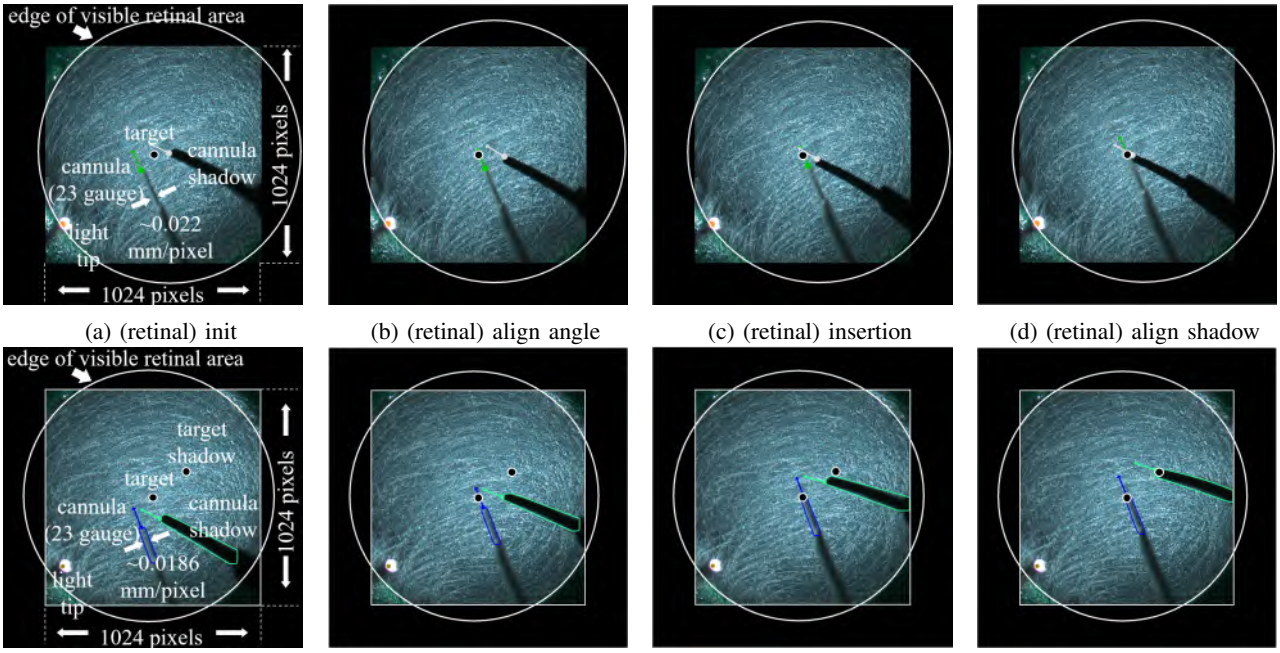
[3] M. Zhou, X. Wang, J. Weiss, A. Eslami, K. Huang, M. Maier, C. P. Lohmann, N. Navab, A. Knoll, and M. A. Nasseri, "Needle localization for robot-assisted subretinal injection based on deep learning," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 8727–8732.

[4] M. Sommersperger, J. Weiss, M. A. Nasseri, P. Gehlbach, I. Iordachita, and N. Navab, "Real-time tool to layer distance estimation for robotic subretinal injection using intraoperative 4d oct," *Biomed. Opt. Express*, vol. 12, no. 2, pp. 1085–1104, Feb 2021. [Online]. Available: https://opg.optica.org/boe/abstract.cfm?URI=boe-12-2-1085

[5] M. Zhou, J. Wu, A. Ebrahimi, N. Patel, Y. Liu, N. Navab, P. Gehlbach, A. Knoll, M. A. Nasseri, and I. Iordachita, "Spotlight-based 3d instrument guidance for autonomous task in robot-assisted retinal surgery," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7750–7757, 2021.

[6] B. Keller, M. Draelos, G. Tang, S. Farsiu, A. N. Kuo, K. Hauser, and J. A. Izatt, "Real-time corneal segmentation and 3d needle tracking in intrasurgical oct," *Biomedical optics express*, vol. 9, no. 6, pp. 2716–2732, 2018.

[7] S. Dehghani, M. Sommersperger, P. Zhang, A. Martin-Gomez, B. Busam, P. Gehlbach, N. Navab, M. A. Nasseri, and I. Iordachita, "Robotic navigation autonomy for subretinal injection via intelligent real-time virtual ioct volume slicing," *CoRR*, vol. abs/2301.07204, 2023. [Online]. Available: https://doi.org/10.48550/arXiv.2301.07204



(a) (retinal) init (b) (retinal) align angle (c) (retinal) insertion (d) (retinal) align shadow

(e) (floating) init (f) (floating) align angle (g) (floating) insertion (h) (floating) align shadow

Fig. 11: The procedure of approaching virtual retinal ((a)-(d)) and floating ((e)-(h)) targets using proposed strategies.

[8] M. Zhou, K. Huang, A. Eslami, H. Roodaki, H. Lin, C. P. Lohmann, A. Knoll, and M. A. Nasseri, "Beveled needle position and pose estimation based on optical coherence tomography in ophthalmic microsurgery," in *2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2017, pp. 308–313.

[9] J. Weiss, N. Rieke, M. A. Nasseri, M. Maier, A. Eslami, and N. Navab, "Fast 5dof needle tracking in ioct," *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, pp. 787–796, 2018.

[10] H. Yu, J.-H. Shen, R. J. Shah, N. Simaan, and K. M. Joos, "Evaluation of microsurgical tasks with oct-guided and/or robot-assisted ophthalmic forceps," *Biomedical optics express*, vol. 6, no. 2, pp. 457–472, 2015.

[11] J. P. Ehlers, A. Uchida, and S. K. Srivastava, "Intraoperative optical coherence tomography-compatible surgical instruments for real-time image-guided ophthalmic surgery," *British Journal of Ophthalmology*, vol. 101, no. 10, pp. 1306–1308, 2017. [Online]. Available: https://bjo.bmj.com/content/101/10/1306

[12] E. M. Tang, M. T. El-Haddad, S. N. Patel, and Y. K. Tao, "Automated instrument-tracking for 4d video-rate imaging of ophthalmic surgical maneuvers," *Biomed. Opt. Express*, vol. 13, no. 3, pp. 1471–1484, Mar 2022. [Online]. Available: https://opg.optica.org/boe/abstract.cfm?URI=boe-13-3-1471

[13] M. Simunovic. Sub-retinal tissue plasminogen activator for sub-macular haemorrhage. Youtube. [Online]. Available: https://www.youtube.com/watch?v=OKYqsVz3Maw&ab_channel=MatthewSimunovic

[14] C. He, E. Yang, N. Patel, A. Ebrahimi, M. Shahbazi, P. Gehlbach, and I. Iordachita, "Automatic light pipe actuating system for bimanual robot-assisted retinal surgery," *IEEE/ASME Transactions on Mechatronics*, vol. 25, no. 6, pp. 2846–2857, 2020.

[15] H. Moon, D. Lee, and D. Nam, "Trimanual technique using assistant-controlled light probe illumination and wide-angle viewing system in 23-gauge sutureless vitrectomy for diabetic tractional retinal detachment," *Ophthalmic surgery, lasers & imaging retina*, vol. 46, pp. 73–6, 01 2015.

[16] T. C. Hutchens, A. Darafsheh, A. Fardad, A. N. Antoszyk, H. S. Ying, V. N. Astratov, and N. M. Fried, "Characterization of novel microsphere chain fiber optic tips for potential use in ophthalmic laser surgery," *Journal of Biomedical Optics*, vol. 17, no. 6, pp. 068 004–068 004, 2012.

[17] B. C. Becker, S. Yang, R. A. MacLachlan, and C. N. Riviere, "Towards vision-based control of a handheld micromanipulator for retinal cannulation in an eyeball phantom," in *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*, 2012, pp. 44–49.

[18] S. Yang, J. N. Martel, J. Louis A Lobes, and C. N. Riviere, "Techniques for robot-aided intraocular surgery using monocular vision," *The International Journal of Robotics Research*, vol. 37, no. 8, pp. 931–952, 2018. [Online]. Available: https://doi.org/10.1177/0278364918778352

[19] S. Yang, R. A. MacLachlan, J. N. Martel, L. A. Lobes, and C. N. Riviere, "Comparative evaluation of handheld robot-aided intraocular laser surgery," *IEEE Transactions on Robotics*, vol. 32, no. 1, pp. 246–251, 2016.

[20] M. Dewan, P. Marayong, A. Okamura, and G. Hager, "Vision-based assistance for ophthalmic micro-surgery," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 49–57.

[21] J. W. Kim, C. He, M. Urias, P. Gehlbach, G. D. Hager, I. Iordachita, and M. Kobilarov, "Autonomously navigating a surgical tool inside the eye by learning from demonstration," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 7351–7357.

[22] J. W. Kim, P. Zhang, P. Gehlbach, I. Iordachita, and M. Kobilarov, "Towards autonomous eye surgery by combining deep imitation learning with optimal control," in *Proceedings of the 2020 Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, J. Kober, F. Ramos, and C. Tomlin, Eds., vol. 155. PMLR, 16–18 Nov 2021, pp. 2347–2358. [Online]. Available: https://proceedings.mlr.press/v155/kim21a.html

[23] P. Zhang, J. W. Kim, P. Gehlbach, I. Iordachita, and M. Kobilarov, "Autonomous needle navigation in retinal microsurgery: Evaluation in ex vivo porcine eyes," 2023. [Online]. Available: https://arxiv.org/abs/2301.11839

[24] Y. Koyama, M. M. Marinho, M. Mitsuishi, and K. Harada, "Autonomous coordinated control of the light guide for positioning in vitreoretinal surgery," *IEEE Transactions on Medical Robotics and Bionics*, vol. 4, no. 1, pp. 156–171, 2022.

[25] Y. Koyama, M. M. Marinho, and K. Harada, "Vitreoretinal surgical robotic system with autonomous orbital manipulation using vector-field inequalities," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2023, pp. 1–7.

[26] M. Zhou, Q. Yu, K. Huang, S. Mahov, A. Eslami, M. Maier, C. P. Lohmann, N. Navab, D. Zapp, A. Knoll, and M. A. Nasseri, "Towards robotic-assisted subretinal injection: A hybrid parallel–serial robot system design and preliminary evaluation," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 8, pp. 6617–6628, 2020.

[27] G. Jocher, A. Chaurasia, and J. Qiu. (2023) Ultralytics yolov8. [Online]. Available: https://github.com/ultralytics/ultralytics