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

Motion Compensation and Adaptive Force Control via iOCT–FBG Sensor Fusion for Robotic Subretinal Injection

Aoqi Long[1,†], Tianle Wu[1,†], Chongyang She[1], Mojtaba Esfandiari[1], Peter Gehlbach[2], Russell H. Taylor[1] and Iulian Iordachita[1]

***Abstract*— Subretinal injection is a highly delicate procedure that demands micron-level precision to avoid irreversible retinal damage. Current robotic systems achieve accurate positioning but remain limited by retinal motion and the lack of tip force feedback. We present the first adaptive tip force compensation framework for robotic subretinal injection, fusing intraoperative optical coherence tomography (iOCT) vision with fiber Bragg grating (FBG) force sensing. Our architecture integrates a finite-state machine (FSM) for surgical phase coordination, a Long Short-Term Memory (LSTM) enhanced residual Kalman filter for real-time motion prediction, and an adaptive compliance estimator for safe force regulation. Compared to previous vision-only and force-only method, *ex vivo* experiments on porcine eyes demonstrate robust improvements: the root-mean-square tracking error reduced by 40% (to 18.5 μm), the maximum absolute error lowered by 2.5 times, and 96.7% of tip forces maintained within ± 0.7 mN. Control delays were minimized to 0.25 s, enabling low-latency corrections beyond freehand capabilities. Our system enhances precision and safety in fragile retinal tissues, advancing the potential for reliable robot-assisted surgeries for retinal diseases.**

I. Introduction

Subretinal injection is a delicate microsurgical maneuver for delivering therapeutic agents between the retinal neuroepithelium and the retinal pigment epithelium (RPE) [1], enabling targeted delivery to photoreceptors and RPE cells. This approach has been successfully applied in gene therapy for Leber congenital amaurosis [2], with ongoing trials for multiple retinal diseases [3]–[5].

Because the retina and RPE are extremely fragile and possess low regenerative capacity, even minor trauma can cause permanent vision loss. Thus, safe subretinal injection requires precise, depth-controlled advancement of a microsurgical insertion through the retinal neuroepithelial layer with accurate tip placement in the potential subretinal space [6]. Freehand procedures are highly challenging due to the surgeon's physiological hand tremor (about 182 μm) [7], the thinness of the retina (depending on location, about 160–290 μm) [8], and limited depth perception using en face viewing systems. Moreover, once the needle enters the subretinal space, the surgeon must simultaneously deliver the agent at a constant, slow rate and hold the tip perfectly still to avoid tissue damage, making sustained freehand control very difficult.

Robotic assistance offers a promising solution. Platforms developed such as [9] and [10] achieve micron-scale positioning accuracy. However, challenges remain, including retinal deformation during tool-tissue interaction and retinal motion caused by respiration and heartbeat. These uncertainties motivate integrating robotic systems with real-time sensing for precise tool positioning and safe interaction.

Intraoperative optical coherence tomography (iOCT) combined with deep neural networks has been used to guide surgical robots by processing intraoperative B- and C-scan data. In this way, accurate tool placement despite tissue deformation is enabled [11]. Feedback-based approaches have demonstrated relatively robust subretinal injections in *ex vivo* porcine eyes [12]. To address retina motion, iOCT-based deep learning methods synchronize robotic tools with tissue dynamics [13], [14]. These approaches can segment surgical needles and retinal layers, especially the internal limiting membrane (ILM) and RPE, which are commonly used as markers of the retina. However, they face limitations such as latency (~ 4 Hz), insufficient training data, poor robustness to unseen conditions, and the absence of force sensing at the tool tip [12].

In terms of force-related safety measures, scleral forces are controlled in previous studies using force-sensing surgical needles equipped with fiber Bragg grating (FBG) sensors, compensating bimanual robots' motion, minimizing scleral forces within a safe limit. [15], [16].

Utilizing the needle tip's contact force to compensate for retinal motion during subretinal injection has not yet been fully investigated. To address this gap, we propose the adaptive tip force compensation scheme for robotic subretinal injection, dynamically estimating tissue compliance to maintain safe interaction forces amid physiological retinal motion. Our robotic system fuses iOCT-based retinal layer segmentation and short-term motion previews with FBG force sensing for post-contact regulation. A residual Kalman filter with Long Short-Term Memory (LSTM) generated previews tracks RPE motion at 200 Hz, while a finite state machine (FSM) coordinates preparation, synchronization, transition, and post-contact phases. Dual controllers blend vision- and force-based velocities through normalized, dwell-adjusted weights, with adaptive compliance tuning to prevent force excursions. The contributions of this work are as follows:

[†]These authors contributed equally to this work.

*This work is supported by the U.S. National Institutes of Health under the grants number R01EB023943, R01EB34397, and partially by JHU internal funds.

[1]A. Long, T. Wu, C. She, M. Esfandiari, R. H. Taylor, and I. Iordachita are with the Laboratory for Computational Sensing and Robotics, Johns Hopkins University, Baltimore, MD, USA.

[2]P. Gehlbach is with Wilmer Eye Institute, Johns Hopkins Hospital, Baltimore, MD, USA.

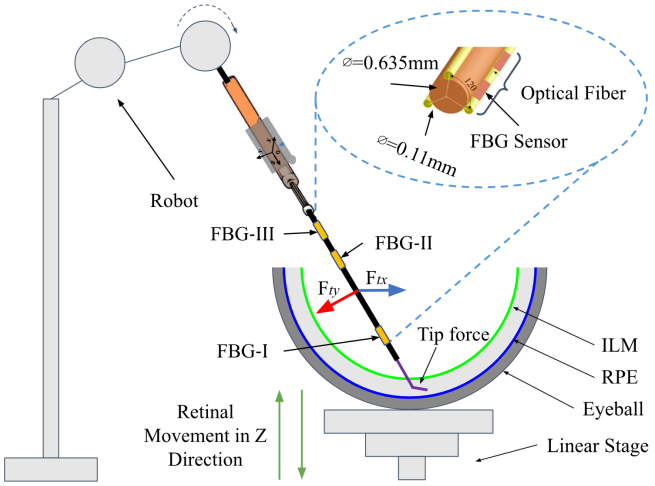

Fig. 1. Device setup for simulating subretinal injection. A cooperative robot holds and actuates an FBG-based force-sensing needle; the eye is mounted on a motorized linear stage that imposes sinusoidal Z-motion to mimic breathing-induced retinal motion. ILM and RPE layers are indicated for context; the inset shows the schematic of the FBG force sensor.

- We developed a unified control architecture—including an FSM for procedural phase management, a residual Kalman filter predictor for bridging low-rate iOCT (4 Hz) with high-rate force feedback (1 kHz), and an adaptive force compensation algorithm that online-estimates tissue compliance to generate smooth reference trajectories and correct deviations.
- We integrated iOCT vision and FBG force sensing with dynamic weight adjustment for velocity blending, reducing tracking RMSE by 40% (to 18.5 μm) under simulated retinal motion and safe tip force regulation (96.7% of samples within ± 0.7 mN) in *ex vivo* porcine eye experiments, while minimizing control delays to 0.25 s.

II. Methods

A. iOCT Vision and FBG Force Sensing

The iOCT B^5-scan, which contains five evenly spaced B-scans over the 4×0.1 mm area, provides retinal depth images at ~ 4 Hz, segmented via a pretrained U-Net [12] to extract needle position, ILM, and RPE in image coordinates. The U-Net was trained on 1707 labeled OCT images using Dice loss and Adam optimizer under the MONAI framework. A lightweight LSTM model forecasts ILM axial motion; a learned layer offset then yields RPE previews ($\hat{z}_{\text{rpe}}$, $\hat{v}_{\text{rpe}}$) in the robot frame via calibrated 1D pixel-to-millimeter mapping as we will discuss in Section II-C.1.

The FBG force-sensing needle monitors tip-tissue interaction at $\sim$ 1 kHz. As depicted in Fig. 1, it comprises three optical fibers recessed in 120° grooves on a 0.63 mm stainless-steel shaft, each with a triplet of 3 mm gratings. The distal triplet (FBG-I) senses intraocular tip bending; proximal triplets (FBG-II/III) compensate for temperature drift and scleral loads similar to [17]. Following [17], the transverse tip force $F_t = [F_{tx}, F_{ty}]^T$ is recovered via a pseudo-inverse calibration matrix, and the controller uses $F_n = \|F_t\|_2$ as a proxy for insertion-axis force, valid under the near-perpendicular approach geometry.

B. Finite State Machine

The FSM (Fig. 2(a)) divides the procedure into four phases that simulating the clinical workflow.

Phase 1 (Preparation): iOCT segmentation identifies needle tip, ILM, and RPE positions. The robot remains stationary while collecting 15 consecutive tip positions. Noise reduction uses interquartile range (IQR) outlier removal followed by median estimation. A 1D transformation maps image pixels p to robot depth Z (mm) with scale $b = 3.379 \times 10^{-4}$ mm/pixel, enabling motion synchronization [14].

Phase 2 (Pre-Contact Synchronization): The robot maintains constant needle-to-RPE distance Δd using vision compensation. A Kalman filter-based predictor estimates current RPE $\hat{z}_{rpe}$ in robot coordinates, yielding desired position: $p_{\text{des}} = \hat{z}_{rpe} + \Delta d$.

A polynomial speed controller computes velocity toward p_{des}, clipped to maximum speed; non-zero user velocity overrides. Transition to Phase 3 occurs if FBG detects a tip force larger than a preset threshold (F_{on}) for a dwell period.

Phase 3 (Transition): Relies on user input until velocity is zero. A 60-sample rolling window tracks force and Δd; upon stop, medians set target force F_{target} and distance, advancing to Phase 4. If force drops below threshold forces (F_{off}) for dwell period, revert to Phase 2.

Phase 4 (Post-Contact): Fuses force and vision with dynamic weights to track motion, prioritizing the high-rate force signal over vision's latency/noise. Surgeon input reverts to Phase 3; low force with positive user speed (retraction) reverts to Phase 2.

The FSM integrates with the robot's velocity-control mode (200 Hz servo loop), allowing seamless surgeon override for shared control.

C. Retinal Position Predictor

The predictor uses a residual Kalman filter (KF) to estimate RPE position and velocity in real-time, bridging iOCT's 4 Hz rate with force's 1 kHz via LSTM previews (Fig. 2(b)).

1) Preview Generation: A pretrained lightweight LSTM network [14] with a single layer of 10 hidden units forecasts the next ILM z-position from the last five historical positions (in pixels); we apply this model autoregressively to obtain two successive predictions (z_1, z_2) for quadratic trajectory fitting. Current z_0, first prediction z_1 (horizon Δt_h), and second z_2 ($2\Delta t_h$) fit quadratic trajectory: $z(t) = at^2 + bt + z_0$ where

$$a = \frac{z_2 - 2z_1 + z_0}{2\Delta t_h^2}, \quad b = \frac{z_1 - z_0}{\Delta t_h} - a\Delta t_h. \tag{1}$$

If implied acceleration $|2a|$ exceeds threshold, fallback to linear ($a = 0$). Discretize into sequences z_{seq}, v_{seq} ($v(t) = 2at + b$); linear interpolation evaluates nominal $z_n(t)$, $v_n(t)$ at time t, with constant-velocity extrapolation beyond horizon.

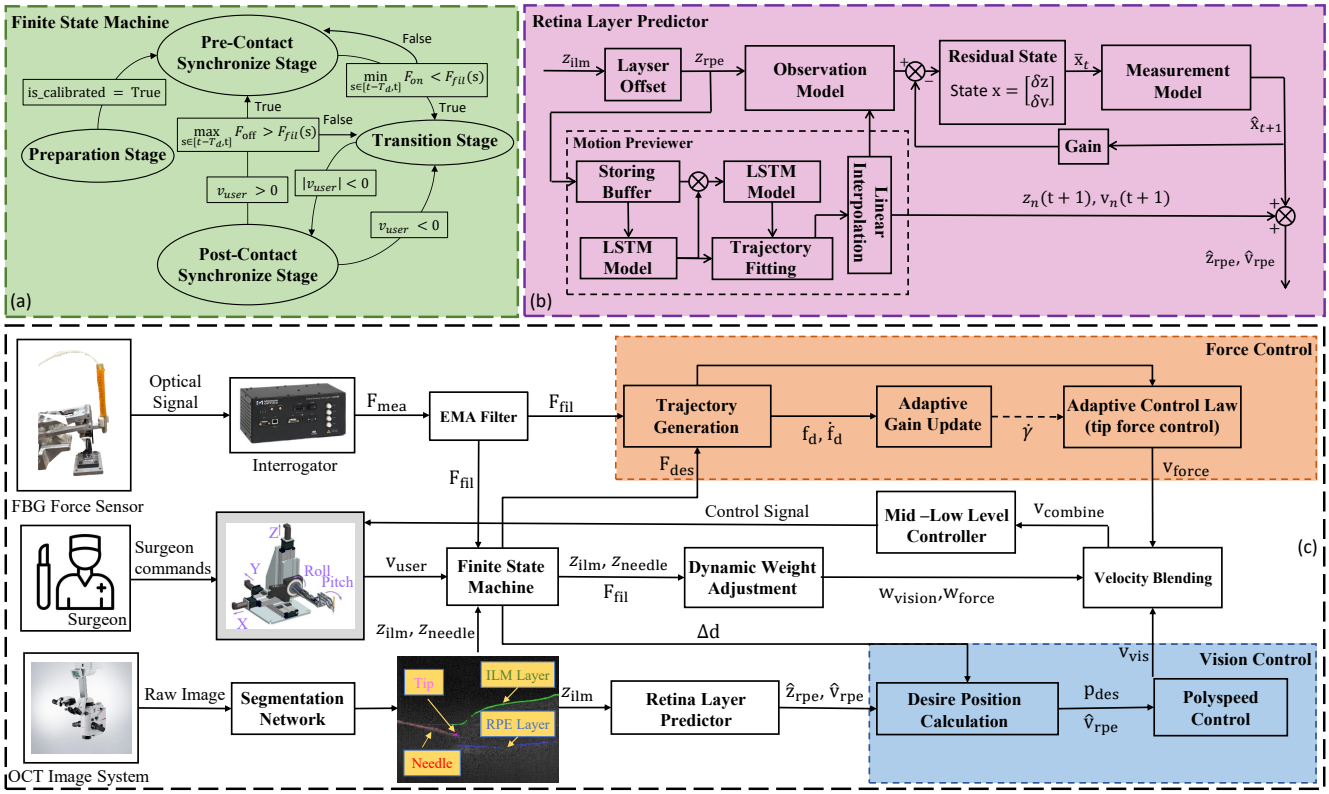

Fig. 2. High-level architecture of the motion compensation system for subretinal injection. (a) A finite state machine divides the procedure into preparation, pre-contact synchronization, transition, and post-contact phases, with transitions based on force thresholds, user velocity, and segmentation confidence. (b) Retina layer predictor generating short-term RPE motion previews via LSTM-based trajectory fitting and refining estimates with a residual Kalman filter updated by high-confidence iOCT frames. (c) Integrated control pipeline fusing iOCT vision (for layer segmentation and motion compensation) with FBG force sensing (for real-time tip interaction regulation), including dynamic weight adjustment for velocity blending, adaptive force control, and polynomial speed control under surgeon oversight.

2) Residual Kalman Filter: The KF tracks residuals $x = [\delta z, \delta v]^T$ relative to preview, where $\delta z, \delta v$ is the position and velocity difference between LSTM prediction and current KF estimation, propagating at control rate and correcting on iOCT arrivals. The state transition model is:

$$\mathbf{x}_k = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \mathbf{x}_{k-1} + \mathbf{w}_{k-1}, \qquad \mathbf{w}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}). \quad (2)$$

where Δt is the time step. The process noise covariance $\mathbf{Q}$ is scaled with Δt and tuned to accommodate variability in breath-like retinal motion. The initial state is set to zero, with a diagonal covariance matrix reflecting uncertainties in position and velocity. The final estimates are computed as $\hat{z}_{rpe} = z_n(t) + \delta z$ for position and $\hat{v}_{rpe} = v_n(t) + \delta v$ for velocity.

3) OCT Update: Upon arrival of an OCT frame, providing the segmented RPE position z_{img} (in mm), segmentation confidence $\text{conf}_{\text{seg}} \in [0, 1]$, and timestamp t_{img}, the residual measurement is computed as $y_{\text{res}} = z_{\text{img}} - z_n(t_{\text{img}})$. The observation model is

$$y = [1, 0]\mathbf{x} + v, \quad v \sim \mathcal{N}(0, R_{\text{img}}) \quad (3)$$

The measurement noise variance R_{img} scales exponentially with staleness and $(1 - \text{conf}_{seg})$ to penalize low-confidence segmentations. The KF performs the correction step using the innovation covariance $S = \mathbf{H}\mathbf{P}\mathbf{H}^\top + R_{\text{img}}$ and Kalman gain $\mathbf{K} = \mathbf{P}\mathbf{H}^\top / S$. The state and covariance are then updated as

$$\mathbf{x} \leftarrow \mathbf{x} + \mathbf{K}(y_{\text{res}} - \mathbf{H}\mathbf{x}), \quad \mathbf{P} \leftarrow (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}. \quad (4)$$

This predictor feeds the robot's high-rate loop, enabling low-latency motion compensation.

D. Adaptive Force Compensation Algorithm

We adopt an adaptive force control scheme based on the estimation of tissue compliance, ensuring that the needle maintains safe interaction forces.

This controller receives raw force data $F_{\text{mea}}(t) = F_n(t)$ from the FBG sensor, filtering it through an exponential moving average (EMA) filter to yield a smoothed signal $F_{\text{fil}}(t)$:

$$F_{\text{fil}}(t) = \alpha F_{\text{mea}}(t) + (1 - \alpha) F_{\text{fil}}(t - \Delta t), \quad (5)$$

where α and Δt are the smoothing coefficient and sampling interval, respectively. The controller then generates a smooth reference trajectory $F_d(t)$ to prevent abrupt loading during injections:

$$F_d(t) = F_{\text{target}} + \Delta F e^{-\lambda(t - t_0)}, \quad \dot{F}_d(t) = -\lambda \Delta F e^{-\lambda(t - t_0)}, \quad (6)$$

where F_{target} is the desired steady-state force learned from phase 3, ΔF defines the initial force offset, λ controls the convergence speed, and t_0 is the time when compensation begins.

The force error is defined as $e_{force}(t) = F_{\text{fil}}(t) - F_d(t)$. If the error exceeds a predefined threshold T, an adaptive control law is triggered. The desired velocity command, generated by the adaptive controller, is defined by:

$$v_{force}(t) = \hat{\gamma}(t)\dot{F}_d(t) - K_f e_{force}(t), \tag{7}$$

where $\hat{\gamma}(t)$ is the adaptive estimation of tissue compliance, and K_f is a proportional gain empirically tuned for stability and responsiveness. The tissue compliance parameter $\hat{\gamma}(t)$ is updated adaptively via:

$$\dot{\hat{\gamma}}(t) = -\Gamma \dot{F}_d(t) e_{\text{force}}(t), \tag{8}$$

where Γ is the adaptation gain, controlling how rapidly the controller adapts to variations in the tissue's mechanical properties. The adaptive force compensation ensures rapid yet stable correction of force deviations, maintaining safe interaction levels during dynamic retinal motion.

E. iOCT Vision Based Control Algorithm

The iOCT-based control computes the desired needle position p_{des}. Velocity toward p_{des} uses proportional control with saturation:

$$v_{\text{vis}} = \min(k_v |p_{\text{des}} - p_{\text{rob}}|, v_{\max}) \cdot \text{sgn}(p_{\text{des}} - p_{\text{rob}}) + \hat{v}_{\text{rpe}} \tag{9}$$

where k_v is the gain, $v_{\max}$ the maximum velocity, p_{rob} the current robot position, and $\hat{v}_{\text{rpe}}$ the predicted RPE velocity for feedforward compensation.

F. Dynamic Weight Adjustment

In Phase 4, the fused velocity is computed as a weighted combination of the vision-based velocity and the force-based velocity, as shown in Fig. 2(c):

$$v_{\text{fused}} = w_{\text{vis}} v_{\text{vis}} + w_{\text{force}} v_{\text{force}}, \tag{10}$$

where v_{vis} is the vision-based velocity, v_{force} is the force-based velocity. Initial weights are $w_{\text{vis}} = 0.3$ and $w_{\text{force}} = 0.7$, since we trust force data more during contact. These weights are dynamically adjusted if certain error conditions persist for longer than the dwell threshold τ_d. Specifically:

- Force error condition: If the absolute difference between the measured force F_{fil} and the target force F_{target} exceeds the force margin, i.e.,

$$e_F = |F_{\text{fil}} - F_{\text{target}}| > F_{\text{margin}} \tag{11}$$

then increment w_{force} by the step size Δw. This adjustment prioritizes the force-based estimate to better align the system with the desired target force.

- RPE stability condition: Experimental observations revealed segmentation errors from the neural network, leading to unstable RPE position. To address this, if the variance of the recent RPE estimates exceeds the threshold, i.e.,

$$\text{Var}(\hat{z}_{rpe}) > \sigma^2_{\text{thresh}}, \tag{12}$$

then increment w_{force} by Δw. This shifts more trust toward the force-based estimate during periods of noisy or unreliable segmentation.

- Positional gap condition: Provided that the segmentation results are stable (i.e., low variance in RPE estimates), if the gap between the robot's position p_{rob} and the RPE estimate falls below the maintained distance Δd minus a safety margin, i.e.,

$$g = p_{\text{rob}} - \hat{z}_{rpe} < \Delta d - g_{\text{margin}} \tag{13}$$

then increment w_{vis} by Δw. This emphasizes the vision-based estimate to avoid over-penetration and maintain the desired separation.

After adjustment, renormalize weights as

$$w_{\text{vis}} \leftarrow \frac{w_{\text{vis}}}{w_{\text{vis}} + w_{\text{force}}}, \quad w_{\text{force}} \leftarrow 1 - w_{\text{vis}}. \tag{14}$$

III. EXPERIMENTS

A. Experiment Setup

The experimental setup, depicted in Fig. 3, comprises a robot developed before [18] configured in velocity-control mode, integrated with a Galil 4088 servo drive, which maintains a joint-velocity control loop operating at 200 Hz. Simultaneously, sensor data acquisition occurs at 1 kHz. The robotic instrument features a custom-designed 33-gauge injection needle embedded with three circumferentially arranged FBG strain sensors of 80 μm diameter each. The sensors interface with an optical interrogator (SM130-700, Micron Optics, USA), enabling precise, real-time measurement of the needle-tip force. Surgical visualization is facilitated by a Leica Proveo 8 microscope, augmented by integrated intra-operative optical coherence tomography (iOCT), with the live B-scan image displayed on an auxiliary monitor (see Fig. 3, left panel).

Experiments are conducted using freshly enucleated *ex vivo* porcine eyes, which are securely mounted on a piezo-actuated linear stage (P-611, Physik Instrumente), capable of replicating physiological retinal motion through precise displacement along the needle-insertion axis. High-level system control, data logging, and visualization tasks are executed using open-source middleware running on a workstation equipped with Ubuntu 20.04.

B. Experimental Procedure

Each experimental trial commences with meticulous mounting and registration of the *ex vivo* porcine eye onto the linear stage, secured within a 3D-printed orbital socket. The linear stage immediately initiates sinusoidal movements with amplitudes of 100 μm at 8 breaths per minute (bpm) to simulate physiological retinal pulsations from the trial outset. The amplitude (0.2 mm peak-to-peak) lies at the center of the 0.1–0.3 mm physiological range reported via iOCT [19], and the frequency of 8 bpm falls within the 6.4–19.5 bpm range observed under patient sedation [20]. Under real-time microscopic and iOCT guidance, the surgeon positions the FBG-instrumented needle tip above the subretinal space via a

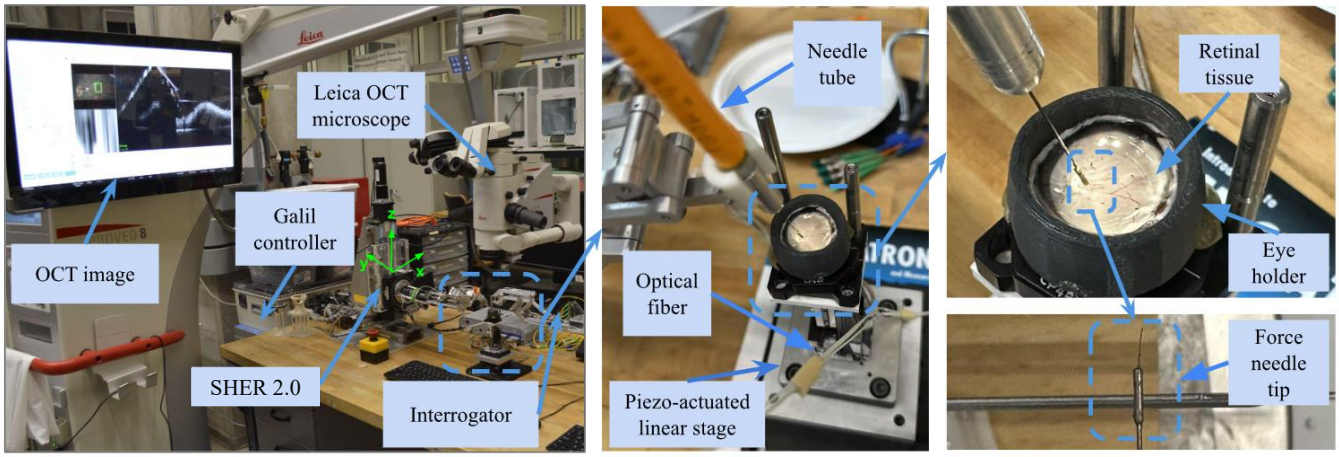

Fig. 3. Experimental setup. A Leica Proveo 8 ophthalmic microscope provided scanning OCT (optical coherence tomography) imaging. A Galil motion controller, interfaced via ROS, commanded the Robot SHER 2.0. Tip forces were measured with a fiber-optic force-sensing needle connected to an optical interrogator. The eye was secured in a 3D-printed holder mounted on a motorized linear stage to impose programmable motion.

pre-established sclerotomy. With data recording (robot joint kinematics, linear stage displacement, iOCT frames, and filtered tip force measurements) initiated immediately after calibration, the system then progresses through the FSM phases detailed in Section II-B:

Phase 1 (Preparation/Calibration): Stationary hold for enabling iOCT segmentation, noise reduction, and LSTM-based motion learning from 15 consecutive frames, with 1D pixel-to-millimeter calibration. Phase 2 (Pre-Contact Synchronization): Maintain constant needle-to-RPE distance ($\Delta d \approx 1\,\text{mm}$) via vision compensation for 5 s. Phase 3 (Transition): Surgeon manually inserts needle into subretinal space until it reaches RPE layers and stops, setting F_{target} and distance medians. Phase 4 (Post-Contact): Activate vision–force fusion for motion compensation, maintaining F_n within $\pm 0.7\,\text{mN}$. Retraction and Re-Synchronization: The surgeon manually retracts the needle, reverting to Phase 2 for 5 seconds to assess re-calibration. Throughout the procedure, the operator always holds a physical emergency-stop button, allowing for immediate termination of robot motion in the event of unsafe behavior.

To evaluate controller robustness and adaptability, we experiment with 10 different *ex vivo* porcine eyes, resulting in a comprehensive dataset of 12 trials. For a subset of porcine eyes, multiple trials are performed at the same retinal site to enable direct, within-site comparisons among the compensation methods. All data are synchronously logged via open-source middleware for subsequent analysis

IV. RESULTS AND DISCUSSION

A. Motion Compensation Performance

Fig. 4 illustrates the trajectories of the needle tip, ILM, and RPE across a representative trial, converted from iOCT pixel coordinates (px) to the robot's local frame (mm) using the calibration factor. The observed trajectories confirm seamless FSM progression: the needle remains stationary during initial registration (Phase 1), achieving stable pre-contact offsets

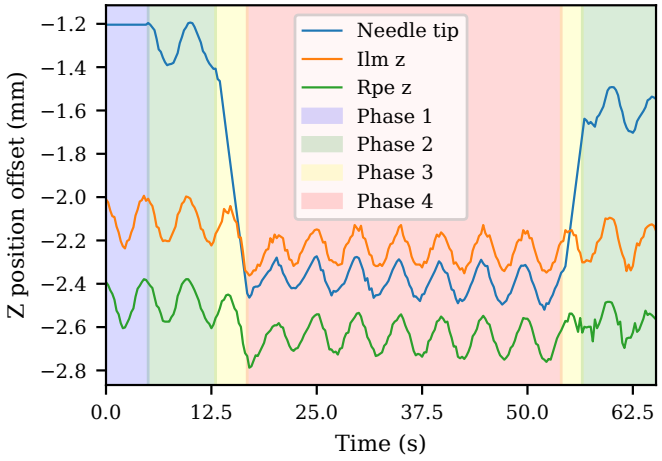

Fig. 4. Trajectories of the needle tip (blue), ILM (orange), and RPE (green) during a representative trial, transformed from iOCT pixel coordinates to the robot's local frame. Shaded regions mark experimental phases: Phase 1 (registration, purple), Phase 2 (pre-contact with motion compensation, green), Phase 3 (insertion, yellow), and Phase 4 (tracking, red).

amid simulated motion (Phase 2). Post-insertion (Phase 3), the fused controller tracks RPE fluctuations with minimal deviation (Phase 4), and retraction enables prompt re-synchronization (back to Phase 2) without loss of registration accuracy.

Table I reports tracking performance across 12 experimental trials using *ex vivo* porcine eyes. In Phase 2 (pre-contact synchronization, relying solely on iOCT vision-based control), we evaluated errors using a known fixed offset (350 px $\approx$ 1.183 mm). In Phase 4 (post-contact regulation, integrating FBG force sensing with vision via dynamic weight adjustment), the offset was estimated robustly as the median of the observed separations, then subtracted so that metrics summarize the residual tracking error $e_t = z_{\text{pred}} - z_{\text{true}}$, in which z_{pred} is the position of the needle with offset, and z_{true} is the position of the RPE layer in OCT coordinate.

We report: Root Mean Square Error (RMSE) as the overall

TABLE I

TRACKING ERROR STATISTICS ACROSS 12 TRIALS. COLUMNS REPORT MEAN, MEDIAN, SAMPLE STD, MIN/MAX, QUARTILES, ROBUST SPREAD (ROBUST $\sigma = \mathrm{IQR}/1.349$), AND CV. CONVERSION: 1 PX $= 3.379\,\mu$M.

Phase	Metric (μm)	Mean	Median	Std	Min	Max	IQR	Q1	Q3	Robust σ	CV (%)
Phase 2	Offset	1182.650	1182.650	0.000	1182.650	1182.650	0.000	1182.650	1182.650	0.000	0.000
	RMSE	31.009	31.127	1.537	28.725	33.290	1.260	30.323	31.584	0.934	4.958
	MAE	26.738	26.501	1.477	24.842	29.249	1.020	26.120	27.140	0.756	5.523
	MaxAE	60.744	60.417	3.670	56.243	66.016	4.798	58.224	63.018	3.557	6.041
	P95AE	53.104	53.729	3.795	47.637	57.686	4.916	50.665	55.581	3.645	7.146
Phase 4	Offset	273.236	266.255	34.652	230.887	320.184	48.158	250.921	299.079	35.699	12.682
	RMSE	18.517	18.993	1.466	16.422	20.240	1.828	17.503	19.331	1.355	7.920
	MAE	15.050	15.300	1.588	12.891	17.152	2.021	13.935	15.956	1.498	10.552
	MaxAE	44.184	43.184	4.568	40.349	53.044	2.352	41.626	43.978	1.743	10.340
	P95AE	35.233	35.175	2.318	31.888	38.838	1.801	34.266	36.067	1.335	6.579

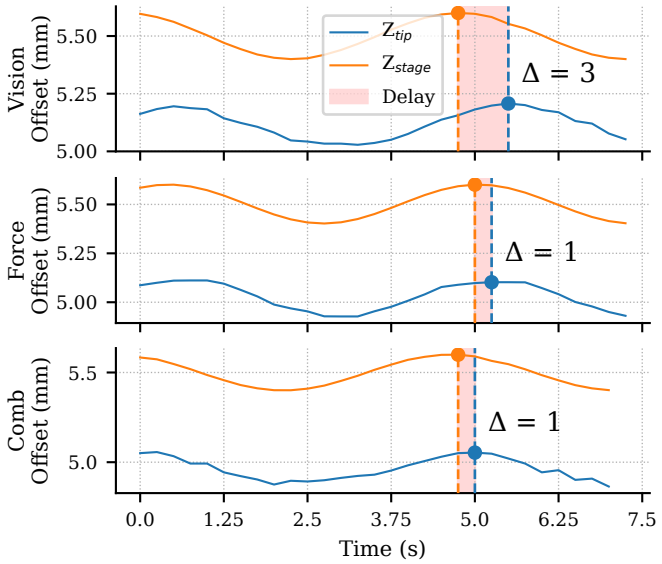

Fig. 5. Example of control delay under three modalities: vision-only (top), force-only (middle), and combined vision–force (bottom). The plots compare retinal motion represented by the linear stage (orange) with the corresponding needle tip position (blue). The shaded red region highlights the temporal offset (Δ) between the two signals, where one sample index corresponds to 0.25 s at a 4 Hz acquisition rate.

performance across different cases; Mean absolute error (MAE) as the typical absolute deviation; Maximum Absolute Error (MaxAE) as the worst case; 95th Percentile Absolute Error (P95AE) as the tail behavior at the 95th percentile; Robust $\sigma = \mathrm{IQR}/1.349$ as the outlier-resistant spread; and Coefficient of Variation (CV %) as the unitless relative variability.

In Phase 2, the vision-only controller achieves a mean RMSE of 31.01 μm and MAE of 26.74 μm, with a MaxAE of 60.74 μm. The P95AE stands at 53.10 μm, indicating reliable but limited precision at larger initial offsets (mean: 1182.65 μm from the RPE). These errors arise from iOCT's 4 Hz update rate, segmentation noise, and uncompensated retinal motion, which can lead to over- or under-estimation of the needle-RPE distance Δd. While sufficient for initial alignment, this phase risks ILM perforation if motion previews deviate significantly.

Phase 4 demonstrates substantial improvements through vision-force fusion, reducing the mean RMSE to 18.52 μm (40% lower than Phase 2) and MAE to 15.05 μm (44% reduction). The MaxAE drops to 44.18 μm (27% improvement), and P95AE to 35.23 μm (34% better), at a closer mean offset of 273.24 μm. The CV for RMSE rises from 4.96% to 7.92%, largely because the smaller offset scale amplifies relative variability, though absolute spread remains tighter (Robust σ: 1.36 µm vs. 0.93 µm). Dynamic weights prioritize the high-rate FBG feedback during contact, mitigating vision latency and ensuring sub-millisecond corrections.

The combined approach benefits from both modalities: vision provides layer previews and depth estimation, while force ensures safe tip-tissue interaction. The 40% RMSE improvement enhances precision, minimizing RPE trauma risks in the fragile subretinal space (160–290 μm thick [8]). Safety is bolstered by adaptive compliance estimation ($\hat{\gamma}(t)$) and dwell-adjusted blending, preventing force spikes during motion. Compared to *ex vivo* benchmarks in the previous work [14], where inside-retina RMSE reached 36.25 μm (MaxAE: 200 μm) without force integration, our fused method halves errors (18.52 μm RMSE), achieving 2.5× better MaxAE control. Robust spread (IQR/1.349) remains low (Phase 4: 1.36 μm), indicating outlier resistance even under segmentation variance ($> \sigma^2_{\text{thresh}}$ triggers w_{force} increase). Overall, the algorithm yields micron-scale precision ($< 20\,\mu$m RMSE), surpassing freehand tremor (182 μm [7]) and prior vision-only systems (36.25 μm RMSE [14]), while force limits ensure safety thresholds.

B. Comparison in Control Delay

Fig.5 illustrates the control delay observed under three different control modalities: vision-only, force-only, and combined vision–force. The plots compare the z-position of the linear stage, which represents retinal motion, and the needle tip position controlled by the robotic end-effector. The temporal offset Δ between the two signals is reported as the difference in sample index, with a sampling rate of approximately 4 Hz (i.e., one sample index corresponds to 0.25 s).

The results show that the vision-only controller exhibits the largest delay, with $\Delta = 3$ (~ 0.75 s). This outcome is

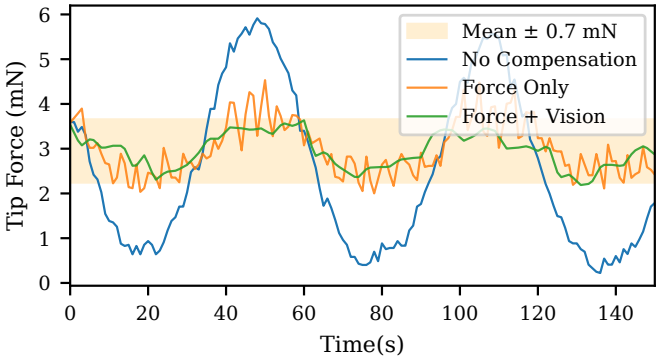

Fig. 6. Comparison of raw tip force data across 140 sample points for three compensation methods: no compensation (blue), force-only (orange), and force + vision (green). The light yellow band indicates the mean force ± 0.7 mN, demonstrating reduced variability with advanced compensation.

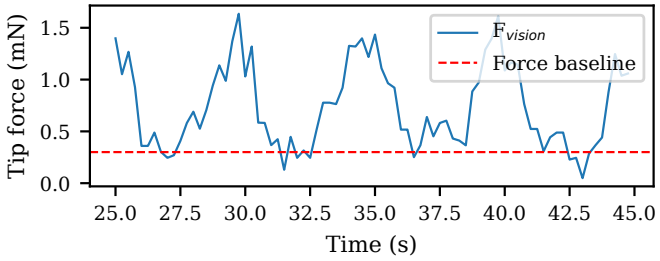

Fig. 7. Contact force at the needle tip under vision-only control. The measured tip force (blue) shows periodic spikes corresponding to retinal motion extrema. The red dashed line indicates the baseline force, representing the minimal contact level when the retina is stationary.

TABLE II

COMPARISON OF FORCE-ONLY VS. FORCE+VISION METHODS ON FORCE VARIATION DURING COMPENSATION. RETINA MOVEMENT IS AROUND $100\,\mu m$ AMPLITUDE, 8 BPM. THRESHOLD $\tau = 0.7\,mN$.

Metric	**Force**	**Force+Vision**
Mean (mN)	1.055	1.142
Median (mN)	0.974	1.106
Std (mN)	0.548	0.296
MaxAE (mN)	1.017	0.929
Proportion within $\pm\tau$ (%)	77.5	96.7
Significance tests (p-values)		
Mann–Whitney U (raw values)	0.103	
Chi-square (within $\pm\tau$)	0.00002**	

**p < 0.001 (statistically significant).

expected, as the vision-based control relies solely on OCT images, which are limited by their acquisition rate of 4 Hz. By contrast, both the force-only and combined vision–force controllers demonstrate substantially reduced delays, with $\Delta = 1$ (~ 0.25 s). This improvement is attributed to the high sampling rate (1 kHz) of the force sensor, enabling faster response to retinal motion.

All data are collected in *ex vivo* porcine eye experiments. The offset measurements are in the robot's local frame for both the linear stage motion and the needle tip position. These results confirm that force sensing substantially reduces control delay. Integrating vision and force yields latency comparable to force-only control while preserving spatial awareness. The force-only modality is *spatially blind*: the FBG sensor measures relative force changes but cannot determine the needle's position relative to ILM or RPE. Consequently, it may maintain nominally correct force while the tip has breached the target layer, or fail to re-establish standoff after retraction. Vision feedback resolves this by continuously providing layer distances, enabling the fused controller to enforce both force and positional safety.

C. Comparison in Force Threshold Control

Fig. 6 presents representative compensation performance assessed via tip force measurements. Experiments were conducted on *ex vivo* porcine eyes using a consistent setup, with sinusoidal retinal motion imposed by a linear translation stage (amplitude 100 μm; frequency 8 bpm) and forces recorded at the needle tip by a force-sensing needle. Under force-only compensation, despite a target constraint of ± 0.7 mN, the tip contact force frequently exceeded this bound. In contrast, the combined force+vision strategy achieved markedly better regulation, with tip forces rarely breaching the threshold. A more detailed numeric performance comparison is shown in Table II.

The combined force–vision method demonstrated improved stability, as reflected in a lower standard deviation (0.296 mN vs. 0.548 mN) and reduced maximum absolute error (0.929 mN vs. 1.017 mN). In addition, the proportion of forces within the safety threshold $\pm\tau = 0.7$ mN increased markedly from 77.5% to 96.7%. Statistical analysis confirmed that while the Mann–Whitney U test did not indicate a significant difference in the overall distribution of force values ($p = 0.103$), the chi-square test showed a highly significant improvement in the proportion of forces within the safe range ($p = 0.00002$). These results indicate that integrating vision feedback with force control enhances both stability and safety by reducing force excursions beyond the clinical threshold.

The vision-only strategy revealed persistent, periodic force spikes at the needle tip. Due to latency in the vision pipeline (Fig. 5), the controller exhibits a ~ 0.75 s lag relative to retinal motion. As the retina reverses direction from its inferior extreme (near the negative displacement peak) and moves upward, the robot continues to advance, driving the needle deeper into the RPE. This over-penetration produces transient force excursions and a quasi-sinusoidal force trace with pronounced peaks near the motion extrema (Fig. 7). Although no consensus "safe" tip force threshold for subretinal injection has been established, we observed several failed insertions associated with excessive forces and damaged RPE layer in Fig. 8.

In our study, force-sensing micro-needles could allow surgeons to detect abnormal resistance in real-time and halt penetration before tissue damage occurs. We attempted to predict the regularity of retinal movement during surgery. However, ocular motion may vary due to physiological factors such as patient anxiety or deep breathing. Because current imaging systems do not provide true real-time feedback, small delays may introduce bias. Future improvements

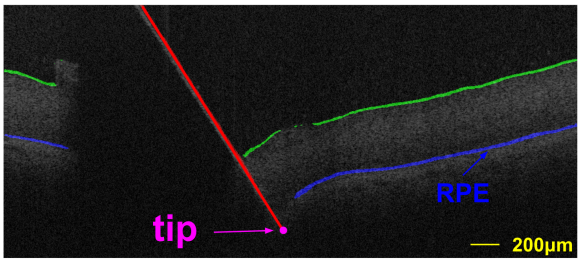

Fig. 8. Example of failed insertion under vision-only control, showing RPE damage due to excessive force. The needle shaft (red line) and tip (pink dot) penetrate the retinal layers.

in image processing speed could enable real-time monitoring and more accurate motion compensation.

All experiments used open-sky *ex vivo* porcine eyes lacking physiological intraocular pressure (IOP), and the single-axis sinusoidal motion omits heartbeat perturbations and irregular breathing [19], [20]; the reported metrics therefore represent a controlled baseline. Future work will extend to closed-globe eyes with physiological IOP, multi-axis motion profiles, and ultimately *in vivo* validation.

V. Conclusion

This work presented an adaptive tip force compensation framework for robotic subretinal injection, integrating iOCT vision and FBG force sensing through a unified architecture comprising an FSM for procedural coordination, an LSTM-enhanced residual Kalman filter for motion prediction, and an adaptive compliance estimator for force regulation. *Ex vivo* experiments on porcine eyes under simulated respiratory motion yielded a 40% reduction in tracking RMSE relative to vision-only control (to 18.5 μm), 2.5$\times$ improvement in maximum error over prior benchmarks [14], 96.7% of tip forces within ± 0.7 mN, and control delays of only 0.25 s. These results show that fusing anatomical context with high-rate force feedback can achieve micron-scale precision and safe tool–tissue interaction in subretinal procedures.

References

[1] Emily Dosmar, Julia Walsh, Michael Doyel, Katlynn Bussett, Adekite Oladipupo, Sabri Amer, and Katherine Goebel. Targeting ocular drug delivery: an examination of local anatomy and current approaches. *Bioengineering*, 9(1):41, 2022.

[2] Stephen Russell, Jean Bennett, Jennifer A Wellman, Daniel C Chung, Zi-Fan Yu, Amy Tillman, Janet Wittes, Julie Pappas, et al. Efficacy and safety of voretigene neparvovec (aav2-hrpe65v2) in patients with rpe65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *The Lancet*, 390(10097):849–860, 2017.

[3] Paul Yang, David Birch, Andreas Lauer, Robert Sisk, Rajiv Anand, Mark E Pennesi, Alessandro Iannaccone, Antonio Yaghy, et al. Subretinal gene therapy drug agtc-501 for xlrp phase 1/2 multicenter study (horizon): 24-month safety and efficacy results. *American journal of ophthalmology*, 271:268–285, 2025.

[4] Wenjing Luo, Lu Guo, Liang Lu, Nan Huang, Yezheng Tao, Yufei Zhang, Yanrong Cao, Shin-Shay Tian, et al. Preclinical studies of an aav8-cyp4v2 gene therapy vgr-r01 for the treatment of bietti crystalline dystrophy. *Molecular Therapy Methods & Clinical Development*, 33(2), 2025.

[5] Muhammad Usman Jamil and Nadia K Waheed. Gene therapy for geographic atrophy in age-related macular degeneration: Current insights. *Eye*, 39(2):274–283, 2025.

[6] Cristina Irigoyen, Asier Amenabar Alonso, Jorge Sanchez-Molina, Maria Rodriguez-Hidalgo, Araceli Lara-Lopez, and Javier Ruiz-Ederra. Subretinal injection techniques for retinal disease: a review. *Journal of clinical medicine*, 11(16):4717, 2022.

[7] Cameron N Riviere and Patrick S Jensen. A study of instrument motion in retinal microsurgery. In *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Cat. No. 00CH37143)*, volume 1, pages 59–60. IEEE, 2000.

[8] Mahnaz Shahidi, Zhangwei Wang, and Ruth Zelkha. Quantitative thickness measurement of retinal layers imaged by optical coherence tomography. *American journal of ophthalmology*, 139(6):1056–1061, 2005.

[9] Botao Zhao, Mojtaba Esfandiari, David E Usevitch, Peter Gehlbach, and Iulian Iordachita. Human-robot interaction in retinal surgery: A comparative study of serial and parallel cooperative robots. In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pages 2359–2365. IEEE, 2023.

[10] Thijs Meenink, Gerrit Naus, Marc de Smet, Maarten Beelen, and Maarten Steinbuch. Robot assistance for micrometer precision in vitreoretinal surgery. *Investigative ophthalmology & visual science*, 54(15):5808–5808, 2013.

[11] Shervin Dehghani, Michael Sommersperger, Peiyao Zhang, Alejandro Martin-Gomez, Benjamin Busam, Peter Gehlbach, Nassir Navab, et al. Robotic navigation autonomy for subretinal injection via intelligent real-time virtual ioct volume slicing. In *IEEE International Conference on Robotics and Automation: ICRA:[proceedings]. IEEE International Conference on Robotics and Automation*, volume 2023, page 4724, 2023.

[12] Demir Arikan, Peiyao Zhang, Michael Sommersperger, Shervin Dehghani, Mojtaba Esfandiari, Russell H Taylor, M Ali Nasseri, Peter Gehlbach, et al. Real-time deformation-aware control for autonomous robotic subretinal injection under ioct guidance. In *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pages 10531–10537. IEEE, 2025.

[13] Demir Arikan, Mojtaba Esfandiari, Peiyao Zhang, Michael Sommersperger, Shervin Dehghani, Russell H Taylor, M Ali Nasseri, Peter Gehlbach, et al. Towards motion compensation in autonomous robotic subretinal injections. In *2025 International Symposium on Medical Robotics (ISMR)*, pages 66–72. IEEE, 2025.

[14] Tianle Wu, Mojtaba Esfandiari, Peiyao Zhang, Russell H Taylor, Peter Gehlbach, and Iulian Iordachita. Deep learning-enhanced robotic subretinal injection with real-time retinal motion compensation. In *2025 IEEE 21st International Conference on Automation Science and Engineering (CASE)*, pages 1285–1291. IEEE, 2025.

[15] Mojtaba Esfandiari, Peter Gehlbach, Russell H Taylor, and Iulian I Iordachita. Bimanual manipulation of steady-hand eye robots with adaptive sclera force control: Cooperative versus teleoperation strategies. *IEEE transactions on human-machine systems*, 2025.

[16] Mojtaba Esfandiari, Ji Woong Kim, Botao Zhao, Golchehr Amirkhani, Muhammad Hadi, Peter Gehlbach, Russell H Taylor, and Iulian Iordachita. Cooperative vs. teleoperation control of the steady hand eye robot with adaptive sclera force control: A comparative study. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 8209–8215. IEEE, 2024.

[17] Xingchi He, Marcin Balicki, Peter Gehlbach, James Handa, Russell Taylor, and Iulian Iordachita. A multi-function force sensing instrument for variable admittance robot control in retinal microsurgery. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1411–1418. IEEE, 2014.

[18] Xingchi He, Daniel Roppenecker, Dominikus Gierlach, Marcin Balicki, Kevin Olds, Peter Gehlbach, James Handa, Russell Taylor, and Iulian Iordachita. Toward clinically applicable steady-hand eye robot for vitreoretinal surgery. In *ASME International Mechanical Engineering Congress and Exposition*, volume 45189, pages 145–153. American Society of Mechanical Engineers, 2012.

[19] Mouloud Ourak, Jonas Smits, Laure Esteveny, Gianni Borghesan, Andy Gijbels, Laurent Schoevaerdts, Yanick Douven, Jesse Scholtes, et al. Combined oct distance and fbg force sensing cannulation needle for retinal vein cannulation: in vivo animal validation. *International journal of computer assisted radiology and surgery*, 14(2):301–309, 2019.

[20] GB Drummond, A Bates, J Mann, and DK Arvind. Characterization of breathing patterns during patient-controlled opioid analgesia. *British journal of anaesthesia*, 111(6):971–978, 2013.