IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

A Novel Hybrid Ureteroscope Tracking for Robotic-Assisted Retrograde Intrarenal Surgery via Recognition of Pathway With Lumen Identification

Jung-Min Han, Dong-Soo Kwon, *Senior Member, IEEE*, and Ki-Uk Kyung, *Member, IEEE*

***Abstract*—Resolving disorientation of the surgeon caused by wrong recognition of scope's position, which often increases procedural time and workload, remains a significant challenge in robotic-assisted retrograde intrarenal surgery (RIRS). This letter introduces a novel hybrid ureteroscope tracking algorithm that integrates low-latency lumen identification with robotic motion data to enhance intrarenal navigation. The system estimates the ureteroscope's position on the centerline of the kidney by recognizing its pathway. In validation tests using a 3D-printed phantom, the proposed method achieved an average localization success rate of 89.2% for major calyx entry and 84.1% for minor calyx entry, with an average computation time of 0.26 seconds, ensuring low-latency operation. Usability testing with ten novice participants demonstrated a 44.5% reduction in cognitive workload (NASA-TLX), improved task success rates, and reduced manipulation effort. These results indicate that the proposed tracking algorithm significantly enhances ureteroscope navigation, improving efficiency and reducing the surgeon's cognitive load in robotic-assisted RIRS.**

***Index Terms*—Surgical robotics: planning, medical robots and systems, vision-based navigation.**

I. Introduction

A. Background

RETROGRADE intrarenal surgery (RIRS) has become the gold standard for kidney stone removal, establishing itself as a widely adopted urological intervention [1]. During the RIRS procedure, precise manipulation of the ureteroscope is essential to prevent tissue injuries and ensure complete stone fragment retrieval [2], [3]. Achieving such dexterity requires surgeons to undergo extensive training, navigating the tortuous and often poorly visible anatomical structures [4]. Although RIRS offers a minimally invasive approach for patients, it imposes significant physical fatigue on surgeons, which can, over time, become a career-limiting burden [5].

To mitigate these challenges, robotic-assisted RIRS has emerged as a promising alternative, introduced by several research groups[6], [7], [8], [9], [10], [11]. These robotic systems, featuring teleoperation capabilities, provide a user-friendly interface designed to alleviate surgeon fatigue [12], [13].

B. Problem Statement

Retrograde Intrarenal Surgery (RIRS) requires precise navigation of the ureteroscope within the kidney. In conventional RIRS(Fig. 2(a)), surgeons rely primarily on video imagery and subtle tactile cues to maintain positional awareness. However, in robotic-assisted procedures, the absence of direct feedback further complicates spatial perception, making it more difficult for surgeons to accurately determine the ureteroscope's depth and orientation. This misjudgment often results in the wrong recognition of the scope's position, causing disorientation of the surgeon during ureteroscopy, leading to prolonged procedural time and increased cognitive workload.

Disorientation is a well-known issue in endoscopic surgery. Studies have shown that minimal access surgery, including endoscopy, poses significant cognitive and spatial awareness challenges for surgeons, often leading to procedural inefficiencies and increased complication risks [14]. To overcome these challenges, an effective ureteroscope tracking system must offer precise and real-time positional awareness while minimizing additional hardware complexity.

C. Related Works

Ureteroscope tracking, which determines the overall position of the ureteroscope within the kidney, offers a potential solution to this problem. Existing tracking methods are broadly categorized into three groups: image-based, external sensor-based, and hybrid approaches [15].

Image-based tracking methods utilize feature or depth information extracted from ureteroscope video imagery [16], [17], [18]. While these methods benefit from simplicity and the lack of additional hardware requirements, they suffer from tracking

Received 4 November 2024; accepted 7 April 2025. Date of publication 28 April 2025; date of current version 16 May 2025. This article was recommended for publication by Associate Editor A. Tzemanaki and Editor J. Burgner-Kahrs upon evaluation of the reviewers' comments. This work was supported in part by R&D Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT under Grant 2022R1A2B5B02002074 and in part by the ROEN Surgical, Daejeon, Korea. *(Corresponding authors: Dong-Soo Kwon; Ki-Uk Kyung.)*

Jung-Min Han is with the Department of the Robotics Program, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea (e-mail: hanjungmin@kaist.ac.kr).

Dong-Soo Kwon is with Roen Surgical Inc., Daejeon 34051, South Korea, and also with the Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea (e-mail: kwonds@kaist.ac.kr).

Ki-Uk Kyung is with the Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, South Korea (e-mail: kyungku@kaist.ac.kr).

This article has supplementary downloadable material available at https://doi.org/10.1109/LRA.2025.3565154, provided by the authors.

Digital Object Identifier 10.1109/LRA.2025.3565154

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

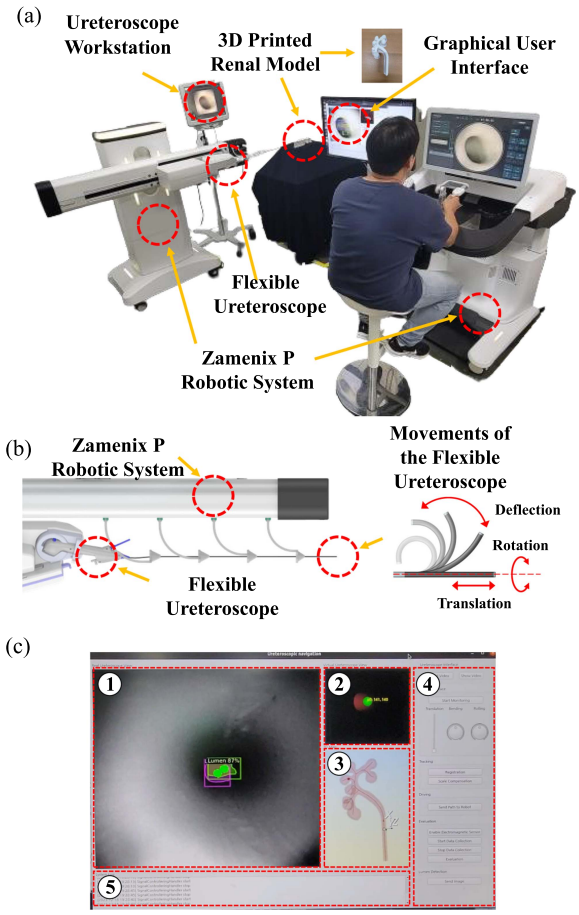

Fig. 1. Experimental setup: (a) Overview of the experimental setup; (b) Ureteroscope movement within the robotic system; (c) Graphical user interface for renal navigation using the proposed algorithm, including ① ureteroscope video feed, ② intrarenal structure rendered in virtual space, ③ navigation view showing the ureteroscope's position, ④ control panel, and ⑤ event log window.

inaccuracies due to image quality degradation caused by intraoperative impurities (e.g., blood clots, dust) and the high structural similarity within the intrarenal environment [15], [21].

External sensor-based methods, using technologies such as electromagnetic (EM) and fiber Bragg grating (FBG) sensors, offer more precise positional data [19], [20], [21]. However, these methods are susceptible to inaccuracies arising from patient movements, such as breathing or coughing, and present challenges in sensor integration with commercial ureteroscopes, potentially compromising device performance.

The hybrid method seeks to combine the strengths of both image-based and sensor-based approaches, achieving continuous, high-precision tracking [22]. However, this method introduces system complexity and requires additional operational space, making it less practical for robotic-assisted RIRS. Addressing these limitations necessitates a system that offers both simplicity and precision.

D. Research Objective

The goal of this letter is to resolve surgeon disorientation caused by errors of the ureteroscope's position, which increases procedural time and workload in robotic-assisted RIRS. To address this, we propose a novel hybrid ureteroscope tracking method grounded in a lumen identification algorithm, as depicted in Fig. 1. Our approach retains the precision of hybrid methods while minimizing spatial and operational complexity by utilizing only robotic motion data and ureteroscope imagery, eliminating the need for external sensors.

Unlike conventional methods, which rely on additional hardware for tracking, our approach estimates the ureteroscope's position along the kidney's centerline, ensuring robust navigation despite minor anatomical variations. This system employs motion data from robotic encoders and centerline information from a 3D ureter model to achieve continuous ureteroscope tracking, as illustrated in Fig. 2(b).

Furthermore, the algorithm identifies lumen bifurcations to guide ureteroscope advancement, enabling the operator to intuitively and continuously ascertain the ureteroscope's position within the ureter, even without additional intraoperative imaging.

II. A Novel Hybrid Ureteroscope Tracking Method

A. The Workflow of the Proposed Method

The workflow of the proposed algorithm is illustrated in Fig. 3. To determine the ureteroscope's position within the ureter, three inputs are used: ureteroscope imagery, robotic motion data, and the centerline of the ureter. Ureteroscope images and robotic motion data are transmitted to the algorithm in real time, allowing it to track the ureteroscope and identify lumens at bifurcation points. The algorithm then maps the ureteroscope's position onto the centerline of the intrarenal structure and renders it within a virtual 3D ureter model for the operator

B. Ureteroscope Tracking Using Motion Data

In kidney stone removal procedures, a thorough inspection of all calyces is critical to ensure a high stone-free rate. To track the ureteroscope within the complex intrarenal structure, the centerline of a 3D ureter model, derived from CT images, was employed. Using a structural analysis toolkit [23], the centerline extending from the ureter to each minor calyx was extracted. This centerline, represented as a graph of points and edges, captures the internal pathways of the kidney and preserves the scale of the original CT model.

For the ureteroscope tracking, we utilized robotic motion data to estimate the ureteroscope's insertion depth rather than relying on external tracking devices, because it provides a direct and stable positional reference, unlike rotation or tip deflection, which require additional sensing mechanisms. As depicted in Fig. 4, the ureteroscope is modeled as a point and vector, and its position is continuously updated based on the translation of the robotic motion data (Fig. 4(b)). When the ureteroscope reaches a bifurcation, the next path is determined by the lumen identification and tracking process.

C. Lumen Identification

Since tracking with motion data alone only estimates the insertion depth, we introduced a method to match lumen positions

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

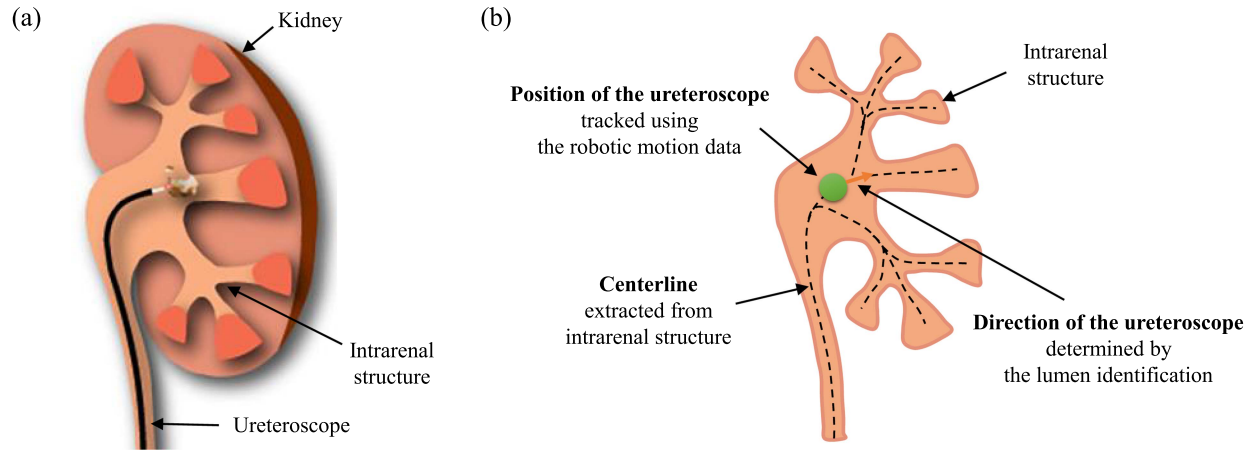

Fig. 2. Illustration of conventional RIRS and the proposed tracking method: (a) Conventional Retrograde Intrarenal surgery (b) the concept diagram of the proposed novel ureteroscope tracking algorithm.

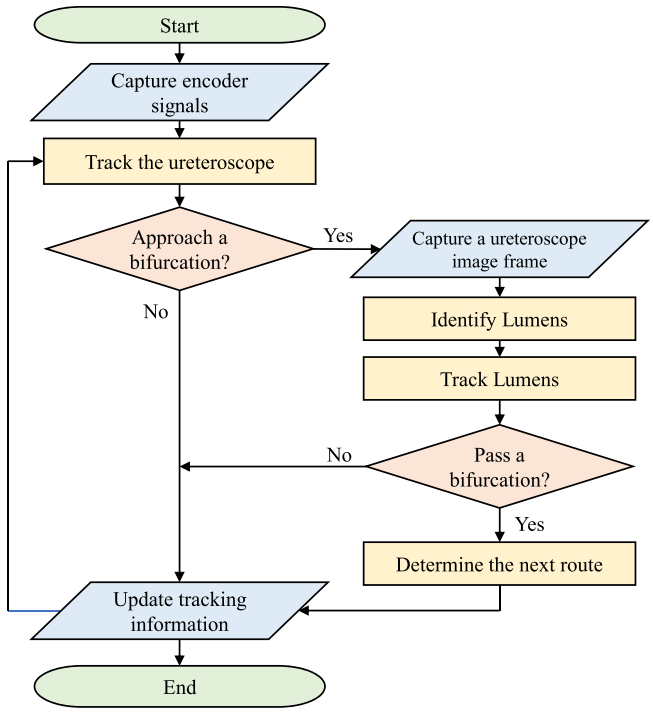

Fig. 3. Workflow of the proposed algorithm.

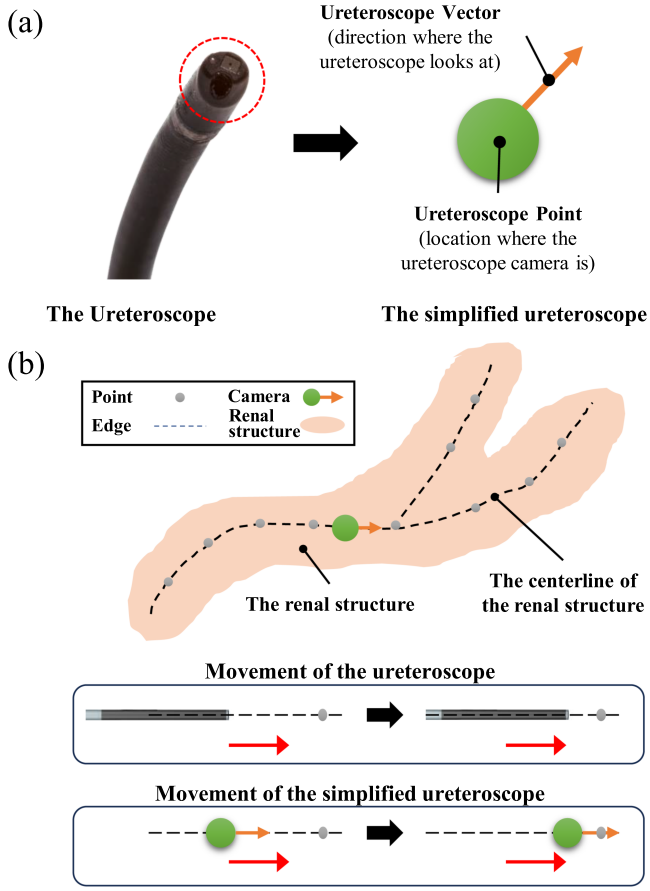

Fig. 4. Ureteroscope tracking along the centerline based on robotic motion data: (a) The distal end of the ureteroscope is simplified as a point and a vector. (b) This simplified model is positioned on the centerline, which is composed of points and edges, and it moves along the centerline as the motion data is continuously provided.

with the intrarenal structure to accurately determine the direction at bifurcation points. This matching process is shown in Fig. 5.

Upon reaching a bifurcation (Fig. 5(a)), points on the next possible paths are selected, and their 3D coordinates are projected onto a 2D image plane (Fig. 5(b)). To ensure accurate projections, the intrinsic and extrinsic camera parameters of the ureteroscope are applied. The projection equation is given by:

$$p' = K_u \times [I|t_{vc}] \times R_{vc} \times P_{vw} \quad (1)$$

Where p' represents the projected point coordinates, K_u is the intrinsic parameter matrix (containing the camera's focal length and image center), and R_{vc} and t_{vc} are the rotation matrix and translation vector, respectively. These transformations map the 3D coordinates from the virtual world coordinate system to the simplified ureteroscope's camera coordinate system, and P_{vw} is the coordinate of the path points on each next path in the virtual world coordinate system (Fig. 5(b)). The intrinsic parameters

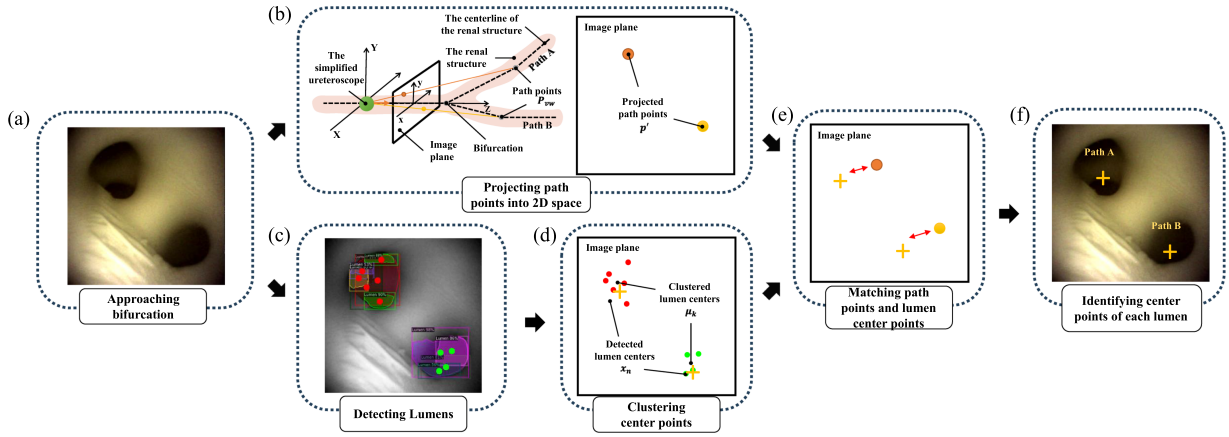

Fig. 5. Lumen identification Process: (a) When the ureteroscope model approaches a bifurcation during navigation, the algorithm is initiated. (b) Path points are projected into 2D space using the ureteroscope's camera parameters. (c) Simultaneously, lumens are detected from the ureteroscope image. (d) The center points of the detected lumens are clustered to determine the lumen center points. (e) A least squares matching algorithm is applied to match the path points with the lumen center points. (f) Finally, the lumens are identified based on the matching results.

were obtained using Zhang's calibration method [24], while the extrinsic parameters were derived from robotic motion data and the point and vector of the simplified ureteroscope model on the centerline.

Simultaneously, we employed instance segmentation using Detectron 2 [25], based on the Mask R-CNN architecture, to detect lumens and extract their center points in the ureteroscope image (Fig. 5(c)). A 3D-printed phantom, reconstructed from patient CT data, was used to capture more than 200 internal ureter images for training the model, incorporating diverse viewing angles. The training process was conducted with a learning rate of 0.001, momentum of 0.9, and batch size of 64 over 900 epochs. The training, conducted on Ubuntu 20.04 with PyTorch 1.8 and CUDA 11.1, took approximately six hours using a system with a Xeon CPU, 128GB RAM, and an RTX 3090 GPU.

Regions of the lumen segmented from a ureteroscope image may be detected separately as multiple individual instances on the image due to the characteristics of the ureteral structure. Therefore, it is necessary to group them through clustering to identify the closest lumen shown in the ureteroscope image. For this purpose, the K-means clustering algorithm proposed by MacQueen [26] was used (Fig. 5(d)). This algorithm is a representative clustering technique that divides data points into K clusters and operates on the principle of minimizing the following objective function.

$$J = \sum_{n=1}^{N} \sum_{k=1}^{K} r_{nk} \|x_n - \mu_k\|^2 \tag{2}$$

Where J is the total distortion, x_n is the detected lumen center point, μ_k is the centroid of the k th cluster which is supposed to be the center coordinates of the identified lumens, and r_{nk} is the variable that is 1 if the point μ_k is the closest point to the point x_n, otherwise 0. The K, the total number of clusters was set equal to the number of projected path points at the bifurcation. Following optimization, the center coordinates of each identified lumen were derived.

Finally, these centers of each identified lumen were then matched with the projected path points using least squares matching (Fig. 5(e)), enabling accurate lumen-path association as the ureteroscope passed bifurcation points (Fig. 5(f)).

D. Lumen Tracking and Direction Decision

Lumen tracking refers to the process of identifying and continuously following the detected lumen structures to assist in navigation, particularly at bifurcation points. Once the lumens are identified, they are continuously tracked to guide the ureteroscope's movement at bifurcation points, as depicted in Fig. 6. Bounding boxes are created around the detected lumen centers to facilitate tracking (Fig. 6(a)). The size of these boxes is selected empirically to balance tracking robustness and the risk of overlap, which could result in tracking loss. A larger bounding box increases robustness but may overlap, while a smaller box risks losing track during rapid ureteroscope movement.

For tracking, we used the Simple Online and Real-Time Tracking (SORT) algorithm, which combines Kalman Filter-based motion prediction with object assignment via the Hungarian Algorithm [27]. SORT enables lightweight, real-time multiple object tracking, ensuring reliable lumen tracking during the procedure. The Euclidean distance between the bounding box center and the ureteroscope image center is calculated for each frame (Fig. 6(b)). Based on these distances, the closest lumen is selected as the next path when the ureteroscope passes a bifurcation (Fig. 6(c)). This process is repeated as the ureteroscope advances through subsequent bifurcations during robotic-assisted RIRS (Fig. 6(d)). If tracking fails due to unexpected occlusions, the system will reset and re-initiate the lumen identification process to re-establish the navigation path.

III. Evaluation of Performance and Usability

A. Experimental Setup

To assess the performance and effectiveness of the proposed method in a robotic setting, an experimental environment was

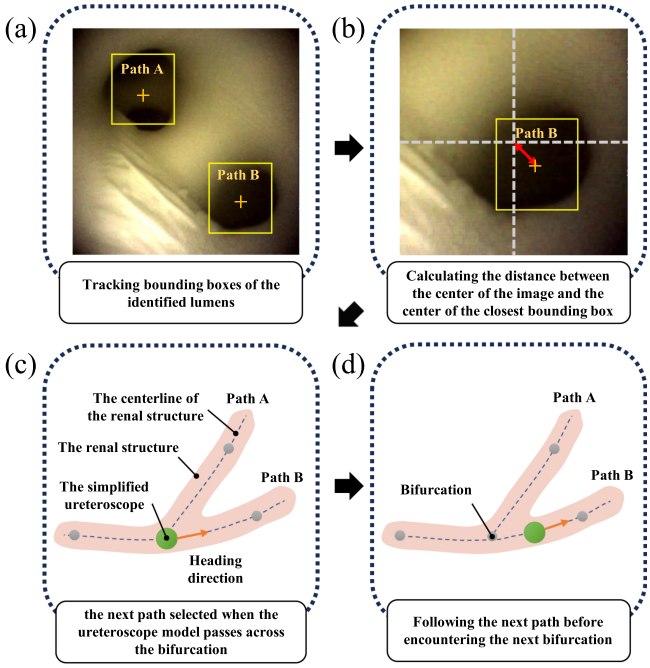

Fig. 6. Lumen tracking and direction decision process: (a) Bounding boxes are generated around the identified lumens and tracked until the ureteroscope model reaches the bifurcation point. (b) The distances between the centers of the bounding boxes and the ureteroscope image are calculated for each frame. (c) Upon passing the bifurcation, the lumen closest to the center of the ureteroscope image is selected as the next path. (d) The ureteroscope model then follows the centerline of the selected path until it encounters the next bifurcation.

created. The method was integrated into a robotic RIRS system as a visual guide to assist the surgeon in tracking the position of the ureteroscope. A disposable ureteroscope, Lithovue (Boston Scientific, US), was mounted on the Zamenix^{TM} P system (ROEN Surgical, Daejeon, Korea) [11], a commercially available robotic platform for RIRS surgery. A 3D-printed renal phantom, created from patient CT data, was accessed using the Navigator ureteral access sheath (Boston Scientific, US). A dedicated software program was developed to display the ureteroscope's position, rendering the 3D kidney model and centerline in real time alongside the master console. To ensure consistent and accurate initialization, a 3D-printed ureter model was fixed at the distal end of the ureteral access sheath. Before tracking began, the robot automatically positioned the ureteroscope at the designated entry point, establishing a reliable starting reference for navigation and enabling precise tracking from the outset. During the procedure, robotic motion data from the Zamenix^{TM} P system and the ureteroscope images from the Lithovue were transmitted to the proposed algorithm, which rendered the tracked ureteroscope's position in the virtual scene.

B. Evaluation of Localization Success Rate and Calculation Time

To evaluate the method's performance, we measured the localization success rate and computation time. Initially, we conducted an experiment using an electromagnetic (EM) sensor attached to the tip of the ureteroscope to obtain a quantitative

TABLE I
LOCALIZATION SUCCESS RATE OF THE PROPOSED ALGORITHM

Pelvicalyceal System		Localization Success Rate
Renal Pelvis		100%
Major Calyx	Upper	82.2%
	Middle	90.0%
	Lower	95.5%
Minor Calyx	#1	93.3%
	#2	86.6%
	#3	100%
	#4	86.6%
	#5	80.0%
	#6	80.0%
	#7	73.3%
	#8	73.3%

TABLE II
COMPUTATIONAL TIME OF THE PROPOSED ALGORITHM

Total time	Computation time for each step		
	Ureteroscope Tracking	Lumen Identification	Lumen Tracking and Direction Decision
0.26s	0.01s	0.24s	0.01s

TABLE III
COMPARISON OF EXISTING METHODS AND THE PROPOSED ALGORITHM

	Tracking Accuracy (mm)	Localization Accuracy (%)	Computation Time (s)
L. O. Maza et al.[16]	1.8	-	0.03
H. Lei et al.[17]	-	76.5	1.26
Z. Huang[18]	15.1	62.5	-
Z. Fu et al.[22]	0.6	-	3.27
Ours	4.13	89.2	0.26

tracking accuracy measurement. The results showed an average tracking error of 4.13mm. While this method provides a quantitative accuracy evaluation, real-time implementation in clinical settings remains challenging due to sensor embedding limitations. As a result, we primarily focused on localization success rate as an alternative performance metric to better reflect the system's robustness in practical scenarios.

For localization success rate, the ureteroscope, mounted on the robot, was inserted into the renal pelvis of the phantom via the ureteral access sheath and navigated from the major calyx to the minor calyx. The actual position of the ureteroscope was compared with the tracked position displayed by the software. This test was repeated fifteen times for each target, and the results are shown in Table I. It was observed that the localization success rate decreased as the number of bifurcations along the path to the minor calyx increased. Nevertheless, the overall average success rate for reaching the major calyces was 89.2%, outperforming previous studies, as shown in Table III. This suggests that our

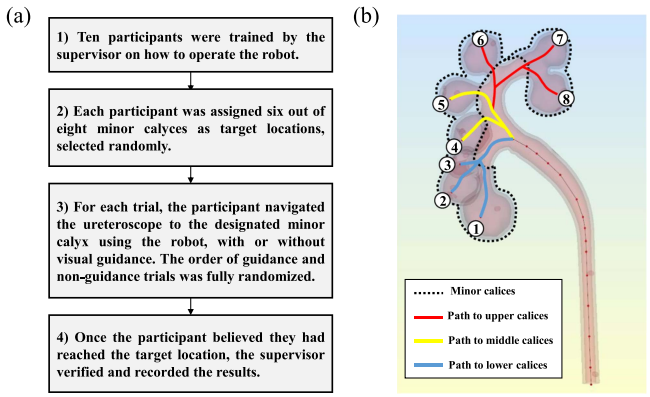

Fig. 7. Usability test with the Zamenix^{TM} P System: (a) Usability test procedure; (b) Location of 8 minor calices in the 3D ureter model.

tracking method enhances ureteroscope navigation accuracy, potentially improving spatial awareness and procedural precision in RIRS.

In terms of computation time, the ureteroscope tracking process updated the position every 0.01 seconds before reaching a bifurcation. At bifurcation points, where lumen identification and tracking occur, the update time was 0.26 seconds, as shown in Table II. These findings indicate that while our method is not the fastest in absolute terms, it achieves a strong balance between accuracy and computational efficiency. This balance of performance ensures efficient guidance without excessive delays, helping surgeons maintain accurate awareness of the ureteroscope's position during tracking.

C. Usability Test Using Zamenix^{TM} P System

A usability test was conducted to evaluate the effectiveness of the visual guidance provided by the proposed method. The procedure is illustrated in Fig. 7(a). Ten novice participants were randomly selected and trained on both the phantom's structure and the operation of the Zamenix^{TM} P system. Each participant was assigned the task of navigating the ureteroscope to three randomly selected minor calyces, chosen from among eight calyces in the 3D model (Fig. 7(b)). To ensure an unbiased comparison, the trials were randomized such that each participant performed three tasks with visual guidance and three without, in a fully shuffled order. Participants were instructed to inform the supervisor when they believed they had reached the target location.

The evaluation metrics for the test included task success rate, cumulative manipulation (translation, rotation, and deflection), and task completion time. Task success rate was defined as the ratio of successful target reach attempts to total attempts. Cumulative manipulation was measured by summing the ureteroscope's motion in terms of translation, rotation, and deflection during task execution. Completion time was recorded from the start of the task until the participant announced that they had reached the target. After the test, the reliability of the results was validated using statistical tools (IBM SPSS Statistics), and the findings are presented in Fig. 8.

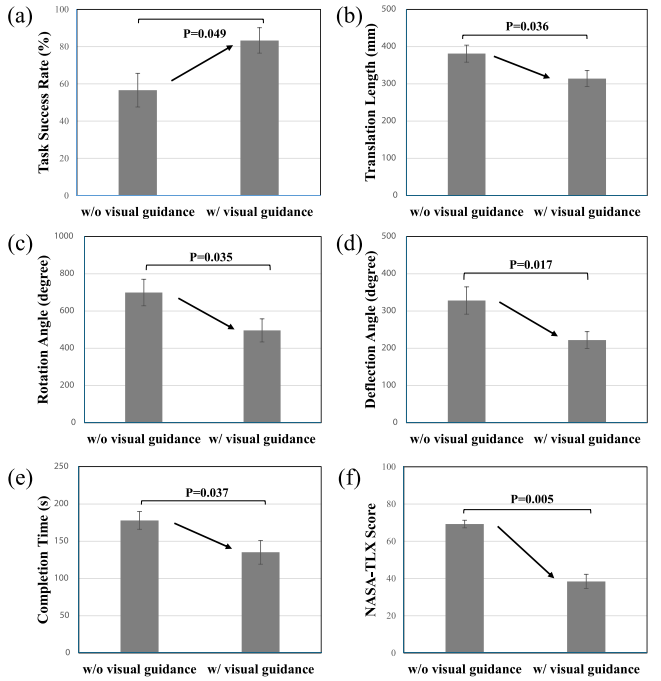

Fig. 8. Usability test results of the proposed algorithm: (a) Task success rate; (b) Cumulative manipulation: translation; (c) Cumulative manipulation: rotation; (d) Cumulative manipulation: deflection; (e) Completion time; (f) NASA-TLX score.

The comparison of performance with and without visual guidance showed that the task success rate improved by 26.6% with the assistance of visual guidance ($p = 0.049$; Chi-squared test). This improvement suggests that the additional spatial awareness provided by the tracking system enables surgeons to locate the target more accurately, increasing procedural reliability.

Regarding cumulative manipulation, the amount of translation decreased significantly by 17.6% ($p = 0.0.049$; Independent Samples t-Test), the amount of rotation decreased significantly by 29.1% ($p = 0.035$; Independent Samples t-Test), and deflection was reduced by 32.2% ($p = 0.017$; Independent Samples t-Test). These reductions indicate that visual guidance helps surgeons optimize ureteroscope movements, leading to more controlled and efficient navigation. By minimizing unnecessary motion, the system could potentially reduce tissue trauma and procedural fatigue, contributing to improved patient safety and surgeon performance.

Furthermore, task completion time was reduced by 24.0% when visual guidance was used ($p = 0.037$; Independent Samples t-Test). Shorter procedure times are clinically significant as they reduce the duration of anesthesia and the potential risks associated with prolonged surgery.

In addition, the NASA-TLX questionnaire [28], administered to assess mental workload, revealed a 44.5% reduction in workload when tracking information was utilized ($p = 0.005$; Wilcoxon signed-rank test). The reduction in cognitive burden suggests that the proposed method allows surgeons to focus more on decision-making rather than struggling with spatial orientation, which could be particularly beneficial for less experienced surgeons.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

These results indicate that the use of the proposed method allowed participants to complete tasks more accurately and in less time, with fewer manipulations required. Additionally, the psychological burden on the operators was significantly reduced, leading to an overall improvement in task efficiency.

IV. Conclusion and Future Work

The goal of this letter is to resolve surgeon disorientation caused by errors of the ureteroscope's position, which increases procedural time and workload in robotic-assisted RIRS. To address this, we propose a hybrid tracking algorithm that integrates lumen identification with robotic motion data, enabling ureteroscope localization without external tracking devices. Unlike conventional methods, which rely on additional hardware for tracking, our approach estimates the ureteroscope's position along the kidney's centerline, ensuring robust navigation even in the presence of minor anatomical variations. Since significant kidney deformation is unlikely during ureteroscopy, the proposed algorithm dynamically adjusts the tracking pathway using lumen identification, ensuring robust navigation despite minor anatomical shifts. The identified lumen positions help determine the appropriate direction of the ureteroscope at bifurcation points, ensuring consistent navigation and reducing the risk of misalignment.

The proposed hybrid tracking algorithm demonstrated promising results, achieving a tracking speed of 0.26 seconds and a localization success rate of 89.2%. This performance was accomplished using only the images from a commercial ureteroscope and the motion data from the robotic system, without the need for any additional tracking devices. This configuration offers several advantages for robotic RIRS. It simplifies the system setup, contributing to more efficient use of space in the operating room. Moreover, because the system only relies on robotic motion data, it is compatible with any commercial ureteroscope, enhancing its versatility.

The results of our comparative experiments showed that the visual guidance provided by the proposed algorithm significantly improved the efficiency of intrarenal navigation. The visualization of the ureteroscope's position, alongside the kidney model, helped the operator better understand both the current position and the location of the target minor calyx. This guidance enabled quicker navigation to the target with reduced manipulation, even in the kidney's complex anatomical structure. Importantly, the tracking information reduced the cognitive burden of determining directions during navigation, allowing operators to perform tasks with greater confidence and less stress.

Thus, we strongly believe that the proposed method overcomes the inherent challenges of identifying the ureteroscope's position within the kidney during robotic RIRS, enhancing both the efficiency and convenience of the procedure.

However, despite its promising performance in a controlled phantom environment, the proposed tracking system has limitations in replicating real kidney conditions. Anatomical variations, tissue elasticity, and the presence of stones or surgical tools may introduce challenges that were not fully addressed in this study. Large stones within the lumen or unexpected deformations in the renal structure could affect lumen identification, leading to tracking inaccuracies. Additionally, the presence of surgical instruments such as laser fibers or baskets may introduce visual occlusions that require further algorithmic refinement. Addressing these challenges will require further investigation in more clinically relevant settings.

To our knowledge, this study is the first to demonstrate the effectiveness of navigation technology for robotic RIRS using a commercial ureteroscope. Its significance lies in the application of our algorithm to a commercial robot and ureteroscope, with successful performance validation, highlighting its potential for real-world clinical use.

In future research, we aim to enhance the system's performance by incorporating techniques such as hardware acceleration and model optimization to reduce its computational complexity. Additionally, we plan to explore the extension of this technology to other endoscopic surgeries, conducting further validation through animal experiments to confirm its applicability and identify areas for improvement within the context of kidney stone removal procedures.

References

[1] T. Inoue, S. Okada, S. Hamamoto, and M. Fujisawa, "Retrograde intrarenal surgery: Past, present, and future," *Invest. Clin. Urol.*, vol. 62, no. 2, pp. 121–135, Mar. 2021, doi: 10.4111/icu.20200526.

[2] B. K. Somani et al., "Pictorial review of tips and tricks for ureteroscopy and stone treatment: An essential guide for urologists from PETRA research consortium," *Transl. Andrology Urol.*, vol. 8, no. Suppl 4, pp. S371–S380, Sep. 2019, doi: 10.21037/tau.2019.06.04.

[3] S. Doizi and O. Traxer, "Flexible ureteroscopy: Technique, tips and tricks," *Urolithiasis*, vol. 46, no. 1, pp. 47–58, Feb. 2018, doi: 10.1007/s00240-017-1030-x.

[4] S. Y. Cho, H. Jeong, M. C. Cho, J. Park, and H. Son, "Current status of minimally invasive surgery for treatment of renal stones and tumors using a flexible ureteroscopy," *J. Korean Med. Assoc.*, vol. 59, no. 6, pp. 459–466, Jun. 2016, doi: 10.5124/jkma.2016.59.6.459.

[5] W. W. Ludwig, G. Lee, J. B. Ziemba, J. S. Ko, and B. R. Matlaga, "Evaluating the ergonomics of flexible ureteroscopy," *J. Endourology*, vol. 31, no. 10, pp. 1062–1066, Oct. 2017, doi: 10.1089/end.2017.0378.

[6] V. Gauhar et al., "Robotic retrograde intrarenal surgery: A journey from 'Back to the Future'," *J. Clin. Med.*, vol. 11, no. 18, Sep. 2022, Art. no. 5488, doi: 10.3390/jcm11185488.

[7] M.-C. Rassweiler-Seyfried, J. Herrmann, J. Klein, M.-S. Michel, J. Rassweiler, and B. Grüne, "Robot-assisted flexible ureterorenoscopy: State of the art in 2022," *Mini-Invasive Surg.*, vol. 6, p. 41, Jul. 2022, doi: 10.20517/2574-1225.2022.41.

[8] R. Saglam et al., "A new robot for flexible ureteroscopy: Development and early clinical results (IDEAL Stage 1-2b)," *Eur. Urol.*, vol. 66, no. 6, pp. 1092–1100, Dec. 2014, doi: 10.1016/j.eururo.2014.06.047.

[9] A. El-Hajj, E. Abou Chawareb, M. Zein, and N. Wahoud, "First prospective clinical assessment of the ILY robotic flexible ureteroscopy platform," *World J. Urol.*, vol. 42, no. 1, Dec. 2024, Art. no. 143, doi: 10.1007/s00345-024-04869-7.

[10] "Robotic-assisted removal of kidney stones for first patient using MONARCH^{TM} platform for urology," *Urol. Times*, Feb. 2023.

[11] D.-H. Lee et al., "ZAMENIX^{TM} R, robotic-assisted retrograde intrarenal surgery system for renal stone removal and its efficacy and safety evaluation," *Front. Biomed. Devices*, vol. 86731, Jun. 2023, Art. no. V0001T07A002, doi: 10.1115/DMD2023-7694.

[12] J. Zhao, S. Wang, J. Wang, J. Li, L. Cui, and J. Li, "Design and experiment of a 3-DoF master device with a 2-DoF parallel mechanism for flexible ureteroscopy," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 19, no. 1, Feb. 2022, Art. no. e2459, doi: 10.1002/rcs.2459.

[13] M. M. Sinha et al., "Technical aspects and clinical outcomes of robotic ureteroscopy: Is it ready for primetime?," *Curr. Urol. Rep.*, vol. 24, no. 8, pp. 391–400, Aug. 2023, doi: 10.1007/s11934-023-01167-4.

[14] C. G. L. Cao and P. Milgram, "Disorientation in minimal access surgery: A case study," in *Proc. Hum. Factors Ergonom. Soc. Annu. Meeting*, 2000, pp. 1806–1810, doi: 10.1177/154193120004402618.
[15] Z. Fu et al., "The future of endoscopic navigation: A review of advanced endoscopic vision technology," *IEEE Access*, vol. 9, pp. 41144–41167, 2021, doi: 10.1109/ACCESS.2021.3065104.
[16] L. Oliva Maza, F. Steidle, J. Klodmann, K. Strobl, and R. Triebel, "An ORB-SLAM3-based approach for surgical navigation in ureteroscopy," *Comput. Methods Biomech. Biomed. Eng.: Imag. Visual.*, vol. 11, no. 4, pp. 1005–1011, Jan. 2023, doi: 10.1080/21681163.2022.2156392.
[17] H. Lei et al., "CT-video matching for retrograde intrarenal surgery based on depth prediction and style transfer," *Appl. Sci.*, vol. 11, no. 20, Oct. 2021, Art. no. 9585, doi: 10.3390/app11209585.
[18] Z. Huang, "Image-matching based navigation system for robotic ureteroscopy in kidney exploration," Ph.D. dissertation, Delft Univ. of Technol., Delft, The Netherlands, 2022.
[19] K. Yoshida et al., "The advantage of a ureteroscopic navigation system with magnetic tracking in comparison with simulated fluoroscopy in a phantom study," *J. Endourology*, vol. 29, no. 9, pp. 1059–1064, Sep. 2015, doi: 10.1089/end.2015.0054.
[20] K. Yoshida et al., "Novel ureteroscopic navigation system with a magnetic tracking device: A preliminary *Ex Vivo* evaluation," *J. Endourology*, vol. 28, no. 9, pp. 1053–1057, Sep. 2014, doi: 10.1089/end.2013.0757.
[21] M. Finocchiaro et al., "Multi-level-assistance robotic platform for navigation in the urinary system: Design and preliminary tests," in *Proc. 11th Joint Workshop New Technol. Comput./Robot Assist. Surg.*, 2022, pp. 90–91.
[22] Z. Fu et al., "Visual-electromagnetic system: A novel fusion-based monocular localization, reconstruction, and measurement for flexible ureteroscopy," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 17, no. 4, Aug. 2021, Art. no. e2274, doi: 10.1002/rcs.2274.
[23] R. Izzo, D. Steinman, S. Manini, and L. Antiga, "The vascular modeling toolkit: A Python library for the analysis of tubular structures in medical images," *J. Open Source Softw.*, vol. 3, no. 25, May 2018, Art. no. 745, doi: 10.21105/joss.00745.
[24] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000, doi: 10.1109/34.888718.
[25] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick, "Detectron2," Accessed: Feb. 13, 2025, [Online]. Available: https://github.com/facebookresearch/detectron2
[26] J. Macqueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probability*, 1967, pp. 281–297.
[27] A. Bewley, Z. Ge, L. Ott, F. Ramos, and B. Upcroft, "Simple online and realtime tracking," in *Proc. IEEE Int. Conf. Image Process.*, 2016, pp. 3464–3468, doi: 10.1109/ICIP.2016.7533003.
[28] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," *Adv. Psychol.*, vol. 52, pp. 139–183, 1988, doi: 10.1016/s0166-4115(08)62386-9.