MEDICAL ROBOTS

AI search, physician removal: Bronchoscopy robot bridges collaboration in foreign body aspiration

Lilu Liu[1]†, Jingyu Zhang[1]†, Fei Wang[1], Jiyu Yu[1], Yuxiang Cui[2], Zhibin Li[3], Jian Hu[4], Rong Xiong[1]*, Haojian Lu[1]*, Yue Wang[1]*

Bronchial foreign body aspiration is a life-threatening condition with a high incidence across diverse populations, requiring urgent diagnosis and treatment. However, the limited availability of skilled practitioners and advanced medical equipment in community clinics and underdeveloped regions underscores the broader challenges in emergency care. Here, we present a cost-effective robotic bronchoscope capable of computed tomography (CT)–free, artificial intelligence (AI)–driven foreign body search and doctor-collaborated removal over long distances via fifth-generation (5G) communication. The system is built around a low-cost (<5000 USD), portable (<2 kilograms) bronchoscope robotic platform equipped with a 3.3-millimeter-diameter catheter and 1-millimeter biopsy forceps designed for safe pulmonary search and foreign body removal. Our AI algorithm, which integrates classical data structures with modern machine learning techniques, enables thorough CT-free lung coverage. The tree structure is leveraged to memorize a compact exploration process and guide the decision-making. Both virtual and physical simulations demonstrate the system's effective autonomous foreign body search, minimizing bronchial wall contact to reduce patient discomfort. In a remote procedure, a physician in Hangzhou successfully retrieved a foreign body from a live pig located 1500 kilometers away in Chengdu using 5G communication, highlighting effective collaboration of AI, robotics, and human experts. We anticipate that this 5G-enabled, low-cost, AI expert–collaborated robotic platform has notable potential to reduce medical disparities, enhance emergency care, improve patient outcomes, decrease physician workload, and streamline medical procedures through the automation of routine tasks.

INTRODUCTION

Foreign body aspiration (FBA) represents a critical emergency medicine issue, affecting individuals across all age groups, particularly children under 5 years old and the elderly (*1*–*3*). If FBA is not addressed promptly, it can lead to severe complications like pneumonia, asthma, or atelectasis (*4*, *5*). Although central airway, namely, trachea, obstructions can often be managed with the Heimlich maneuver (*6*, *7*), around 80% of foreign bodies (FBs) lodge in the bronchi, where bronchoscopy remains the gold standard for both diagnosis and treatment (*8*, *9*). Effective bronchoscopic intervention typically requires preoperative computed tomography (CT) scans to precisely locate, size, and orient the foreign object, alongside highly skilled practitioners to perform the search and removal (*10*). However, such resources are frequently unavailable in community clinics and underdeveloped regions, highlighting a broader challenge in emergency care—where timely intervention is critical but often unattainable because of resource limitations (*11*, *12*).

The rapid development of medical robotics has transformed complex surgeries across multiple disciplines, enhancing precision and improving patient outcomes (*13*–*16*). Systems like the da Vinci Surgical System (*17*, *18*) have extended surgeons' capabilities in urology, gynecology, and general surgery by offering superior precision, flexibility, and control. Similarly, the MAKO robotic system (*19*, *20*) aids orthopedic surgeons in preoperative planning and intraoperative guidance for knee and hip replacements, resulting in improved alignment and longevity of implants. In neurosurgery, platforms such as ROSA (*21*) and Neuromate (*22*) provide real-time three-dimensional (3D) imaging and navigation, enabling surgeons to operate on delicate brain and spine structures with minimal invasiveness. Despite these advances, current robotic systems are often bulky and expensive and primarily serve as surgical assistants rather than autonomous collaborators, limiting their utility in emergency situations, especially in resource-constrained settings. In recent years, enhancing autonomy has become a focus in medical robotics, and several autonomous robotic systems have demonstrated effectiveness in in vivo settings. For example, for laparoscopic surgery, Saeidi *et al.* (*15*) introduced the Smart Tissue Autonomous Robot, designed for autonomous anastomosis without physician intervention. For tissue biopsy, Kuntz *et al.* (*23*) developed an autonomous needle-steering robot for biopsy, integrating a CT-based planning and electromagnetic positioning system to increase the autonomy. For bronchial FBA surgery, an ideal solution would allow physicians to perform multiple procedures remotely, with reduced response times and physician workloads on each one. In such a scenario, minimal onsite personnel would be needed, with the robot autonomously performing most of the procedure and requiring physician intervention only at critical stages, enhancing emergency care by bridging the human and robot expertise (*24*, *25*).

Achieving this vision requires medical robots capable of autonomously performing certain tasks, relieving physicians from managing every step of the procedure (*26*, *27*). In bronchoscopy, the FBA surgery necessitates a thorough search process, namely, exploration, for all possible bronchi to localize the FB, which is a time-consuming and skill-intensive task with a high risk of secondary injury to patients, especially when an external navigation method, such as preoperative CT imaging and planning, is unavailable. If robots could

[1]Department of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China. [2]Zhejiang Humanoid Robot Innovation Center, Ningbo 315000, China. [3]Department of Computer Science, University College London, London WC1E 6BT, UK. [4]Department of Thoracic Surgery, First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310009, China.
*Corresponding author. Email: ywang24@zju.edu.cn (Y.W.); luhaojian@zju.edu.cn (H.L.); rxiong@zju.edu.cn (R.X.).
†These authors contributed equally to this work.

autonomously locate and safely navigate to the object, physician intervention could be reserved for the final removal stage. However, current robotic bronchoscopy systems, such as the Ion (*28*, *29*) and Monarch platforms (*30*, *31*), rely heavily on preoperative CT imaging or external positioning systems because of the lack of an intelligent system with long-term memory that organizes the on-the-fly perception into spatial awareness like a human expert, which is crucial for efficient bronchial exploration.

Exploration is a fundamental prerequisite for robot applications in unknown environments. The exploration process consists of three parts: perception, path planning, and control. Perception involves mapping and localization using cameras or other external sensors, path planning generates trajectories for exploration, and control ensures precise movement execution. In bronchoscopy, exploration methods are categorized into camera-only and CT-based approaches. Camera-only methods, such as manual bronchoscopy, are widely used in clinical practice but require expertise and pose risks of disorientation (*32*). Simultaneous localization and mapping (SLAM)–based techniques offer a potential solution by building sparse maps through image feature matching, as seen in robotics applications (*33*, *34*). However, in medical contexts like colonoscopy (*35*), laparoscopy (*36*), and bronchoscopy (*37*), SLAM suffers from low-quality mapping, limited field of view, and nonrigid organ motion, making clinical deployment challenging. CT-based methods reconstruct high-quality dense bronchial maps preoperatively, allowing preplanned navigation to target FBs or lesions. Intraoperative localization often relies on external sensors, such as electromagnetic tracker and fiberoptic shape sensors (*38*), but their sensitivity to respiratory motion introduces inevitable localization errors. To mitigate this, some methods use image-based registration between real and CT-rendered images (*39*–*44*), although CT-to-body visual divergence leads to localization inaccuracies. Consequently, most CT-based methods assist doctors in guidance rather than enabling autonomous control of a robot. Although some studies have investigated semiautonomous control in a bronchoscope robot, continuous human intervention remains necessary (*45*, *46*).

Here, we present a fifth-generation (5G)–based remote robotic system for bronchoscopy, integrating artificial intelligence (AI)–driven autonomous search and enabling doctor-robot collaboration for FBA surgery (Fig. 1 and Movie 1). The sensor of our robot is only a camera, without the need for any preoperative CT or external positioning system. Our system leverages AI algorithms to sense, think, and act, where a tree structure is used as a key role for long-term memory, allowing for an online-built compact topological map of the bronchial tree from an ever-growing bronchoscopic image stream. On the topological map, a tree-search planner and a policy neural network jointly make the search decision, enabling thorough and safe lung coverage and robot navigation for FB search. The hardware platform, designed for affordability and portability, includes a mobile base, an electric slide, and a bronchoscope robot weighing ~2 kg. The bronchoscope features a 3.3-mm-diameter catheter with a 1.2-mm working channel, a miniature endoscopic camera, and LED (light-emitting diode) illumination. The removal system uses two linear motors to control forward motion and the actuation of 1-mm biopsy forceps. The entire system costs less than 5000 USD, making it accessible to community clinics and in underdeveloped regions. Both virtual and physical simulations on realistic human bronchial models validated the system's ability to autonomously navigate the lung, replicating the performance of an experienced physician. In vivo experiments further demonstrated the feasibility of remote operation: A doctor in Hangzhou successfully removed an FB from a pig located 1500 km away in Chengdu via 5G communication, with the physician only intervening during the final removal stage. This successful trial across AI, robotics, and human expertise underscores the potential of this low-cost, AI expert–collaborated robotic platform to address medical resource disparities, improve emergency care, enhance patient outcomes, reduce physician workload, and streamline medical procedures by automating routine tasks.

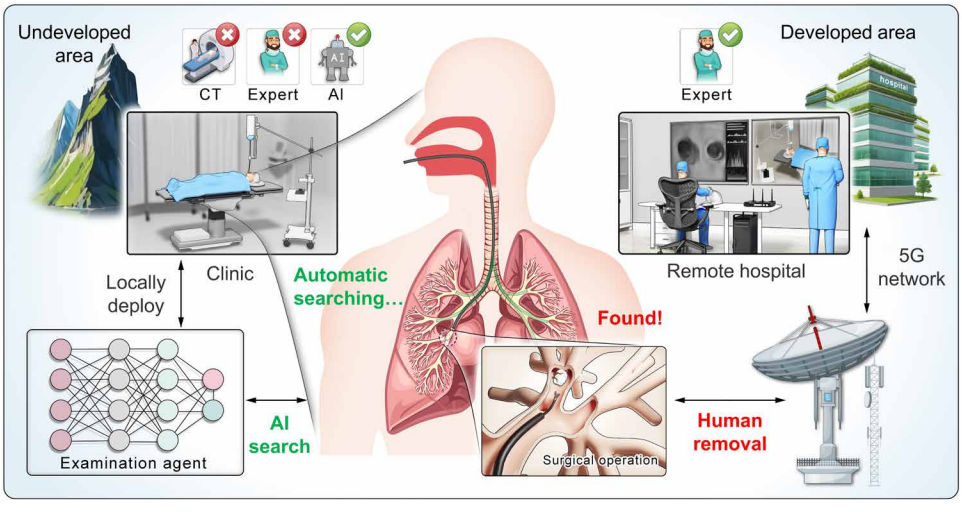
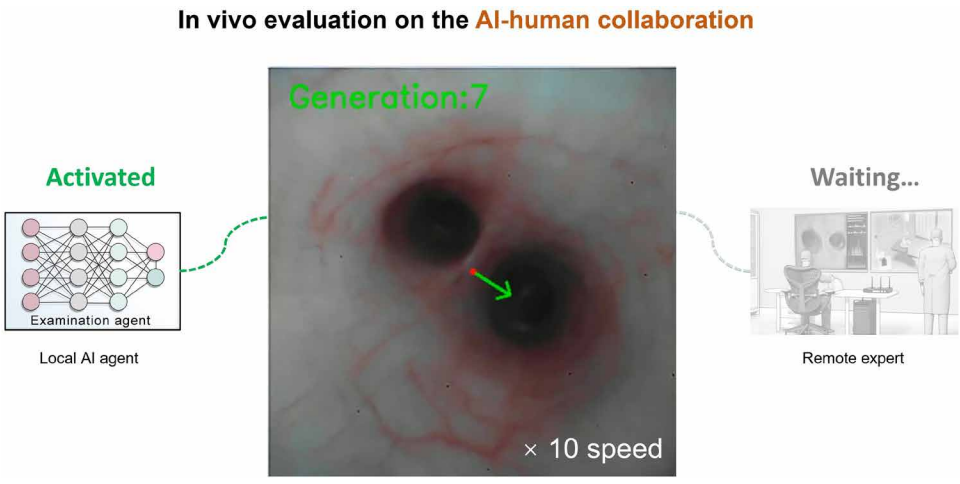


Fig. 1. Illustration of the AI-human collaboration remote bronchoscope robot system. Our system is designed to address FBA in undeveloped areas, where access to CT imaging and experts is limited. The system consists of two key components: a low-cost bronchoscope robot located in undeveloped areas and an AI agent for safe and automatic searching for FBs and collaboration with a remote expert via a 5G network. Upon presentation of a patient with suspected FBA, the AI agent autonomously navigates the bronchoscope to identify the FB within the bronchus. Once detected, the system alerts a remote specialist, allowing them to take control of the bronchoscope robot for FB removal.

RESULTS

AI-human collaboration for bronchial FBA surgery

The procedure was divided into five distinct stages. In the first stage, local clinic personnel inserted the bronchoscope and activated the AI agent. During the second stage, the AI autonomously conducted a full bronchial examination, navigating on the basis of the neural network strategy while the supervising physician remotely observed the process. In the third stage, upon identification of the FB, the physician could override AI control to halt the bronchoscope's advancement, allowing clinic staff to introduce the biopsy forceps into the working channel. The fourth stage involved the physician remotely operating the forceps to grasp the FB. In the fifth stage, the AI resumed



Movie 1. Overview of the AI-human collaboration remote bronchoscope robot system for FBA surgery.

control, extracting the bronchoscope along with the FB from the bronchi. The entire process of the procedure is presented in movie S1. Throughout this process, only the fourth stage necessitated direct physician involvement, minimizing the overall demand on the expert's attention and cognitive load. The AI autonomously handled the more time-intensive tasks in stages two and five, substantially reducing the workload on the physician. Conversely, stages one and three required only basic technical skills, easily managed by clinic personnel. This division of labor maximized efficiency, assigning tasks according to respective expertise and ensuring optimal patient outcomes.

AI agent design and learning strategy for bronchial FB search

The AI agent featured two core components for FB search: One was the policy neural network for visual perception and robot control, and the other was the tree-like memory bank (TLMB) for long-term memory storage and exploration planning (Fig. 2A). The input to the policy neural network comprised a sequence of recent images captured from the catheter's camera, along with the planned exploration guidance from the TLMB. The network's output included the FB probability, the airway bifurcation probability, and the safe steering action. The FB probability indicated whether the FB was found, and if so, the remote expert would be alerted for FB removal. The airway bifurcation probability indicated whether the robot was located at airway bifurcation, and if so, the latest image would be processed and stored in the TLMB. The safe steering action was used to control the catheter tip to follow the guidance and remain centered in the airway. To account for the effects of respiration in real-world scenarios, image sequences of the recent bronchial generation were buffered and sampled as the input of the policy network, avoiding the redundant identification at the same bifurcation caused by the coupling of respiratory motion and robot motion.

Considering the natural tree structure of the bronchial anatomy, the TLMB was designed to build a highly compact tree-like topological map where the tree nodes and edges could efficiently store all historical memory during the FB search process. Correspondingly, a depth-first search (DFS)–based planner was used to generate the exploration guidance and guided the policy network for a thorough search. The input of the TLMB was triggered as if the predicted bifurcation probability of the policy network was larger than 0.5. In this situation, the latest image was treated as the bifurcation image and processed by a lumen detection network (fig. S14). When the bifurcation image with detected lumens was input into the TLMB, the topology map created a new node to store the bifurcation image and built edges according to detected lumens as potential branches, namely, frontiers, to be explored. At each bifurcation, the DFS-based planner output the stop command, determined whether the robot should advance into a new branch, or returned to a previously explored one on the basis of the current tree node where the robot lay and unexplored frontiers after DFS. If advancing into a new branch, the DFS-based planner fed the policy neural network an exploration guidance to control the robot, and the current tree node expanded to create leaf nodes representing unexplored branches. If returning, the planner switched to a previous parent node as the exploration guidance. As the tree structure grew, the memory dynamically incorporated new perception, allowing for the planner and controller to increasingly search the whole bronchial tree until the FB was found.

For training, a 3D airway model segmented from the CT scan was used to establish the simulated environment (see Supplementary Methods for details), allowing real-time rendering of bronchoscopic images. This setup enabled the training of the policy neural network under the guidance of the tree structure to search all bronchial branches. Three key components during training were the expert agent, domain randomization, and Sim2Real adaptation (Fig. 2B). The expert agent provided privileged information in the simulated environment, generating exploration guidance on the basis of the robot's location and previously explored regions, as well as ground-truth bifurcation labels and steering actions for the policy network supervision. The domain randomization component randomly generated the cyclic motion of respiration and placed FBs in different bronchial branches for better generalization. The Sim2Real adaptation compensated for the visual differences between rendered and real bronchoscopic images, using adversarial learning to ensure that the policy network could generalize to real-world scenarios with zero-shot transfer.

Simulation results and in vitro evaluation of bronchial exploration

We first assessed the effect of sampling on the image sequence input in the simulation model (fig. S11). Our results indicated that sampling substantially improved the coverage of exploration of the policy neural network (Fig. 3, A to C, and fig. S15). Important metrics used in this study are defined in Supplementary Methods. Compared with the group inputting a continuous image sequence with a fixed length, the sampling group, which uniformly sampled frames with a fixed length from historical images as input, could explore more branches (103.5 ± 3.5 versus 46.0 ± 1.0) within the same time. The key factor in this performance boost was that the sampling strategy effectively simulated the coupling between respiration and

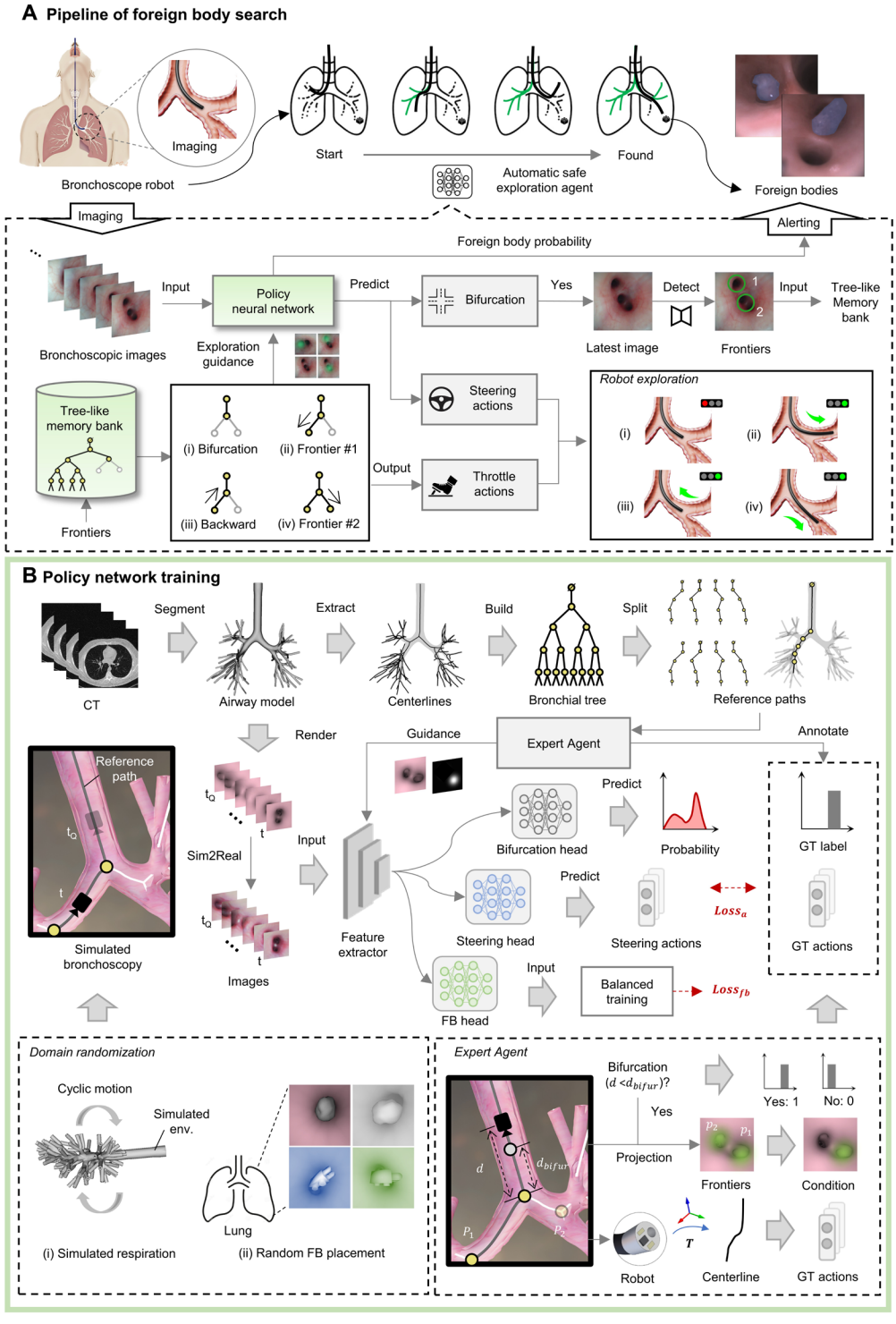


Fig. 2. Pipeline of the AI agent and training strategy for an FB search. (**A**) Overview of the AI agent for an FB search via bronchial exploration. The AI agent is composed of two hierarchical components for an FB search: a policy neural network for visual perception and control and a TLMB for long-term memory and exploration planning. The policy network processes recent images from the catheter's camera along with TLMB guidance, generating predictions for the FB presence, airway bifurcation detection, and safe steering actions. The TLMB constructs a topological tree map of the bronchial anatomy, enabling efficient memory storage and exploration planning. A DFS-based planner generates exploration guidance, controlling the robot's movement to ensure a thorough bronchial search. (**B**) Training strategy of the policy neural network. The policy neural network comprises a feature extractor and three task-specific heads for foreign object identification, bifurcation recognition, and steering action prediction. The feature extractor is based on a convolutional neural network, and these task-specific heads are based on MLPs. During the training stage, a CT scan is used to establish the simulated 3D airway environment through segmentation of the airway model. Subsequently, the bronchial tree is built by extracting centerlines from this model. These centerlines are divided into subsets to serve as reference paths. An expert agent is introduced to automatically generate guidance and annotations for the policy neural network on the basis of the reference paths and the simulated robot position. To address the sparse distribution of FBs, the FB detection head is trained using balanced training. The neural network's input consists of image sequences rendered from the simulated environment, capturing the view from the last bifurcation t_Q to the current position t. To enhance network generalization, Sim2Real adaptation and two domain randomization techniques are used, enabling robust transfer to real-world scenarios for an efficient and accurate FB search.

robot motion. This enhanced the neural network's robustness to variations in image sequences caused by the coupled motion, a critical factor when addressing real-world scenarios where respiratory and robot motions interact. This strategy was expected to improve the system's resilience to the influence of respiratory motion on input image sequences in clinical settings. Next, we evaluated the lumen detection performance at airway bifurcations by incorporating depth estimation into the network (Fig. 3D and fig. S14). Given the large depth variation in the airway, the integration of depth prediction with lumen detection provided the network with both geometric and appearance cues, leading to enhanced accuracy (93.97% versus 91.38% when the threshold was 16 pixels).

In the in vitro model, we allowed the AI to autonomously complete the search of all bronchial branches, incorporating simulated breathing into the phantom. A human expert with more than 20 years of experience and more than 200 cases per year and who was trained over 100 trials of teleoperation was invited to perform the same task

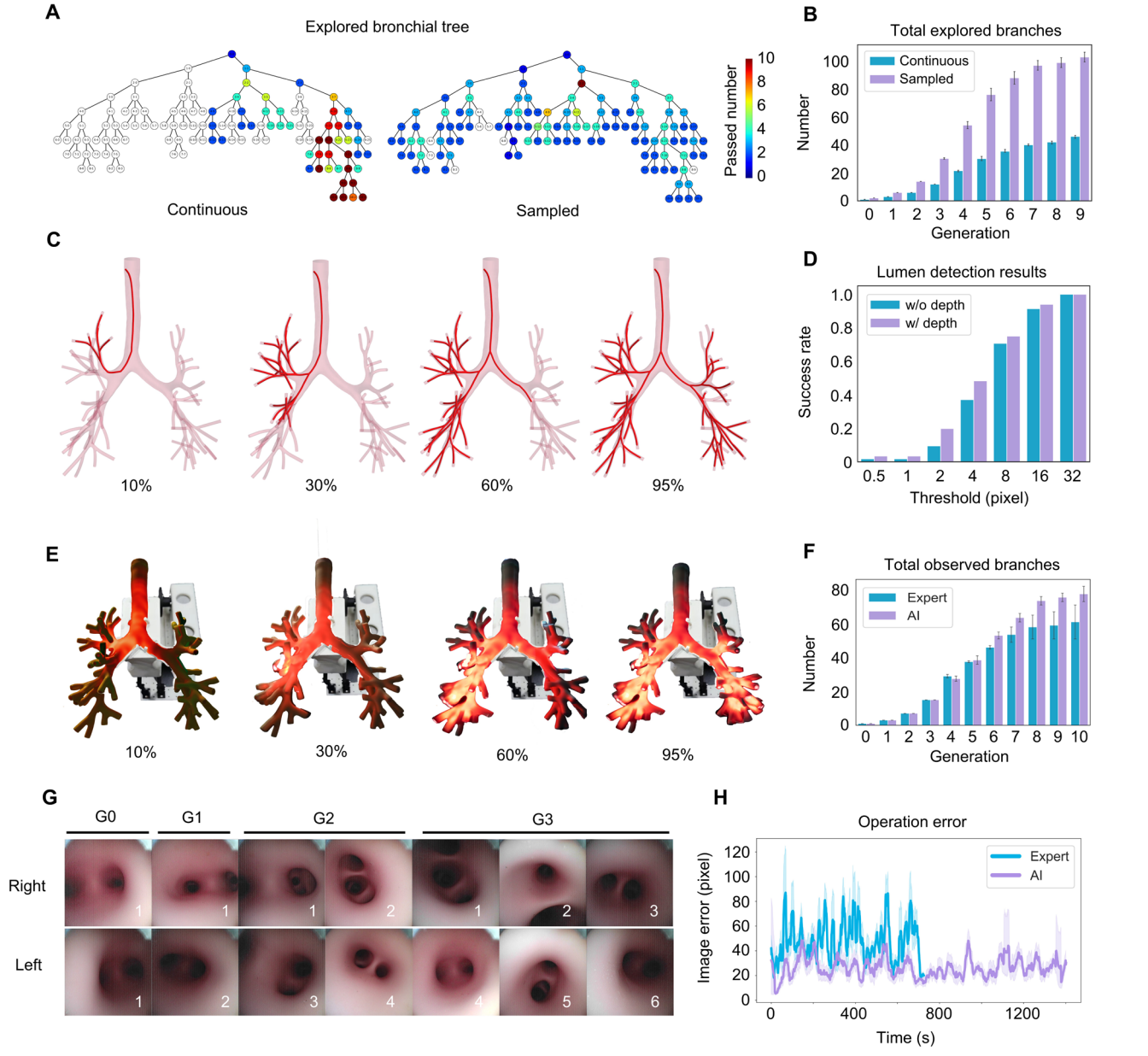


Fig. 3. Exploration results of in experiments. (**A**) Results of AI agent exploration in the simulation environment. "Continuous" refers to the input of 10 consecutive frames to the policy network, whereas "Sampled" indicates the input of 10 historically sampled frames. (**B**) Exploration coverage. With the increase in the bronchial generation (going deeper), the number of explored branches also increases. The Sampled method achieves higher coverage compared with the Continuous method. Bar plots with error bars depict the mean values ± SDs, with $n = 5$ independent experiments. (**C**) Exploration trajectory. The red path shows the robot's trajectory during bronchial exploration, where the policy network safely and comprehensively covers the entire bronchial tree. (**D**) Performance of lumen detection. We evaluated two networks: one learning both depth and detection and another learning detection only. (**E**) Exploration process in the phantom. The illuminated path indicates the robot's trajectory during bronchial exploration in the phantom. (**F**) In vitro exploration results. We recorded the total observed branches with the increase in bronchial generation. Bar plots with error bars show the mean values ± SD, with $n = 3$ independent experiments. (**G**) Observed bronchoscopic images in the phantom. G0 to G3 indicate bronchial generations from the trachea (generation 0) to the third-generation bronchi. The numbers overlaid on the images represent the identities of different branches within the same bronchial generation. Both the left and right bronchi are included. (**H**) Operational errors during in vitro exploration. During the exploration process, all image frames were captured to evaluate image errors and to compare the operational accuracy of the AI against that of human experts. The numbers of recorded frames were $n = 21{,}800$ for the expert group and $n = 22{,}578$ for the AI group. Line plots with error bands show time-varying mean image errors ± 95% confidence intervals (CIs) within a time window of 10 frames during each exploration process. The operational error of the AI agent was notably lower than that of the human expert.

as the AI agent for comparison. The results, shown in Fig. 3F, indicated that the AI consistently outperformed the human expert, particularly as the AI navigated deeper into the bronchial branches (number of total observed branches of 61.0 ± 9.8 versus 77.5 ± 10.2 at the 10th generation). This was expected given the exponential increase in the number of leaf nodes in the tree structure as the system explored deeper bronchi. Fatigue and cognitive load could cause human operators to lose track of their progress, leading to incomplete exploration, a challenge the AI effectively overcomes. This suggested that the AI could serve as a more effective search tool compared with a human expert in complex bronchoscopy tasks. Furthermore, the policy network exhibited better central retention, as evidenced by its training in simulation, resulting in lower deviations from the airway centerline compared with manual exploration (28.45 ± 0.57 versus 17.67 ± 0.54 pixels of operation error). This reduced deviation was likely to minimize potential airway wall damage, indicating that the AI agent provided superior patient safety. Figure 3E illustrates our process of exploring the entire bronchial tree with the robot, captured through long exposure, highlighting the DFS as reflected in the lit bronchial branches. This visual representation effectively grounded the classical tree search algorithm in the context of bronchial exploration.

Because of the passive compliance of the catheter's transition section, unavoidable multiple contact points between the catheter and the bronchial wall occurred during movement. To measure this, a contact force testing experiment was conducted, as detailed in Supplementary Methods. Results showed that the contact force of the catheter's body with the bronchial wall was less than 0.4 N when steering the catheter's tip without contact, which is relatively small in bronchoscopy and acceptable in clinical practice. Moreover, in bronchoscopy procedures, scopes are recommended to be parallel to the axis of the airway to avoid perforating the airway or breaching a vascular structure. It is also critical to maintain the proper spatial orientation and a clear view of the working field unobscured by blood or mucous (*47*, *48*), demonstrating the benefits of following the centerline of the bronchus during bronchoscopy.

In vitro and in vivo demonstration of remote bronchial FB removal

We conducted telesurgery experiments using an in vitro phantom, selecting a remote site in Chengdu, ~1500 km from Hangzhou. Figure 4C illustrates the 5G network performance at this site, with an average download bandwidth of 211.41 ± 5.48 Mbps (megabits per second) and an upload bandwidth of 88.29 ± 3.95 Mbps, sufficient for high-quality image transmission and remote control. As shown in Fig. 4D, although the transmission delay was higher with the transmission control protocol (TCP) than the user datagram protocol (UDP), the frame loss rate on the 5G network remained extremely low, and the policy was robust against frame loss, making UDP ideal for transmitting images for remote control. Figure 4E highlights the network's load capacity, where the delay slightly increased with higher image resolutions. At resolutions greater than 450 pixels by 450 pixels, packet loss became noticeable, although the overall latency (around 40 ms) remained well within the human reaction time range (200 to 300 ms), confirming minimal influence on the human operator's performance. Details of the 5G communication framework are described in Supplementary Methods.

To further validate the system, we conducted remote FB removal experiments using a breathing phantom. Four different FBs (Fig. 4B) were randomly placed into various branches (0 to 6) of the bronchial model (Fig. 4A), with each group repeated 10 times. The results (Fig. 4, F to H) showed no statistically significant difference between remote and local operations for FB removal ($P > 0.05$). However, Fig. 4I demonstrated that the shape of the FB significantly influenced retrieval success rates ($P < 0.05$). In our study, a safety guarantee mechanism was designed to monitor the real-time state of the network connection. When the network connection was lost for more than 5 s, the forceps were automatically withdrawn, and the local AI was activated to safely control the bronchoscope robot back to the trachea (airway of generation 0), waiting for further judgment by the physician and clinic staff, as shown in movie S2.

For in vivo testing, we deployed the bronchoscopy system in a remote clinic in Chengdu, with the teleoperation platform based in a hospital in Hangzhou, 1500 km away (Fig. 5, A and B). The in vivo experiment used a 5-month-old Bama miniature pig purchased from Chengdu Dossy Experimental Animals Co. Ltd. It was 1.3 m long, weighed 38.5 kg, and was in good health (Fig. 5C). For comparison, we performed a CT scan, AI-guided bronchoscopy exploration, and expert bronchoscopy exploration on the pig. Figure 5D presents the pulmonary airway model reconstructed from the CT scan. Images captured during the AI-guided exploration are shown in Fig. 5E. Although the CT scan provided superior coverage of peripheral bronchi (<3 mm), it was less effective for detecting small, low-density FBs. In contrast, bronchoscopy remained the gold standard for FBA surgery in the main and lobar bronchi (>3 mm). We emphasize that the CT scan here was only used for demonstration and comparison and was not used by the robot or AI in any stage.

Comparing AI-guided and expert bronchoscopy (Fig. 5, F and G), AI showed comparable coverage of exploration in the lower-level airways (less than the fourth generation) but notably outperformed the expert in navigating deeper bronchi (more than the fifth generation). As depicted in Fig. 5H, the AI-controlled robot maintained better central alignment than manual operation (30.44 ± 0.24 versus 20.37 ± 0.30 pixels), resulting in less airway wall contact and reaching deeper airway levels faster (Fig. 5G and movie S3). Specifically, the AI-controlled robot reached the seventh generation within 300 s and the 10th within 340 s, using only 67.02% of the time used by the expert to reach the same generation and 71.51% to reach the eighth generation. As depicted in Fig. 5I, the AI-guided navigation group examined a substantially larger number of branches than expert bronchoscopy (43.25 ± 2.25 versus 53.0 ± 2.0). In the FB retrieval trials, eight experiments were conducted with a 100% success rate in locating and retrieving FBs. Figure 5J presents images of the retrieved FBs, demonstrating the system's efficacy in search and clamping, with expert intervention completing the retrieval.

Compared with CT-guided FBA surgery, AI-driven bronchoscopy reduced procedure times. As shown in Fig. 5K, the CT scanning, FB identification, path planning, and bronchoscopy procedures took a total of 13.31 ± 2.88, 14.24 ± 6.74, 12.58 ± 2.91, and 1.88 ± 0.18 min, respectively. In total, the CT-guided surgery took on average ~42 min. In contrast, AI-driven surgery only focused on the bronchoscopy procedure, including FB search and removal, which took 4.72 ± 2.00 minutes in total (movies S4 to S7). The experimental results highlighted the higher efficiency of our AI-driven method (movie S8). These results confirmed the system's robust sim-to-real transfer capability, with zero-shot in vivo performance closely matching the in vivo performance observed in the phantom model.

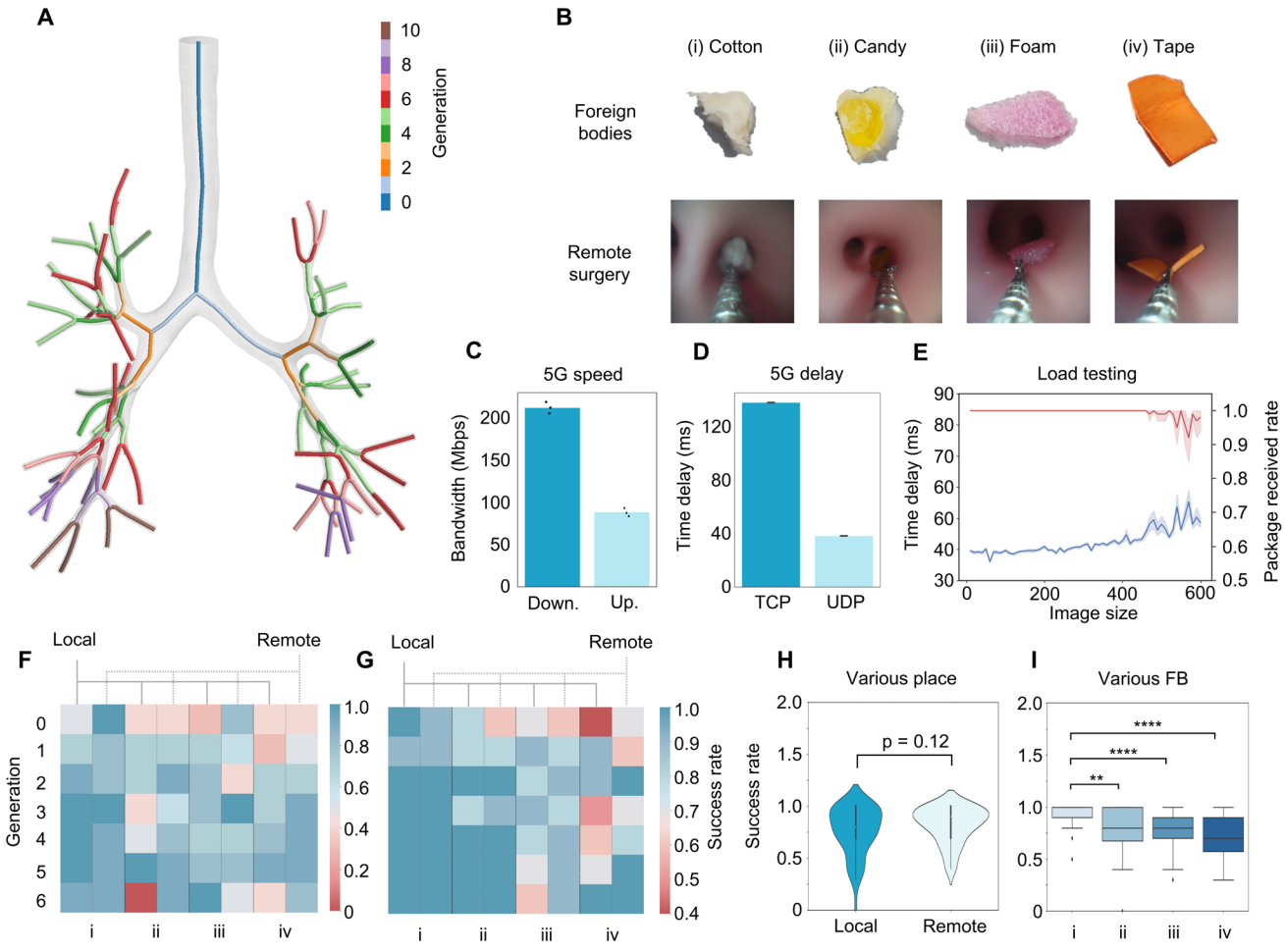


Fig. 4. Results of in vitro remote control experiments. (**A**) Representation of the bronchus model and bronchial generations. (**B**) Types of FBs and corresponding grasping images. (**C**) Upload and download bandwidth of the 5G network. (**D**) TCP and UDP communication latency over the 5G network. Bar plots with error bars show the mean values ± SD, with $n = 379$ independent experiments. (**E**) Load testing of the 5G network. A total of 60 image sizes, ranging from 0 pixels by 0 pixels to 600 pixels by 600 pixels at 10-pixel intervals, were chosen. For each image size, $n = 100$ independent experiments were conducted to evaluate the time delay and package received rate. Line plots with error bands depict mean values ± 95% CIs. (**F**) Grasping of FBs under respiratory motion. In both remote and local experimental settings, four types of FBs were placed at 0 to 6 bronchial generations, with each group tested 10 times to calculate success rates. (**G**) Grasping of FBs without respiratory motion. (**H**) Distribution of grasping success rates in remote versus local settings, showing no significant difference ($P > 0.05$). (**I**) Success rate distribution for the four types of FBs, indicating significant differences between FBs. For each type of FB, the number of independent experiments for grasping was $n = 280$. Box plots show the median, interquartile range (IQR), and data range within 1.5× IQR, with outliers. The paired *t* tests were conducted to show statistical significance for each comparison of different groups. $^{**}P < 0.01$ and $^{****}P < 0.0001$.

DISCUSSION

Here, we present the decoupled modules for planning and control in the AI system of the low-cost bronchoscope robots. This system autonomously executes the entire process of bronchial FBA surgery without the need for preoperative CT scans or external positioning systems. Furthermore, when taken over by the remote physician, the system easily switches to the control only with the guidance given by the physician. The success of our approach hinges on two core components. The first is the hierarchical architecture decoupling planning and control, where two decision-making modules are operated at different frequencies across different decision horizons, allowing for any time autonomous operation or physician intervention. The second is the application of sampling techniques, virtual experts, and style transfer methods to reduce the gap between virtual simulations and real-world conditions, enabling effective zero-shot transfer from simulated to practical environments.

At the heart of our system lies a tree-structured topological map, which plays the role of long-term memory bridging the visual perception with the safe robot control and exploration planning. This structure not only enhances the interpretability of decision-making processes but also meets the interactive requirements of medical professionals, addressing limitations commonly associated with black-box neural networks. By leveraging prior anatomical knowledge of the bronchus, the tree structure organizes redundant and online-acquired image streams into a compact, structured topological memory, notably improving spatial awareness. Moreover, our experiments demonstrate that neural networks perform exceptionally well in controlling the catheter's movements during respiratory actions using an end-to-end control strategy. This approach effectively overcomes the challenges posed by real-time perception and modeling within complex bronchial environments, eliminating the need for explicitly designed control strategies.

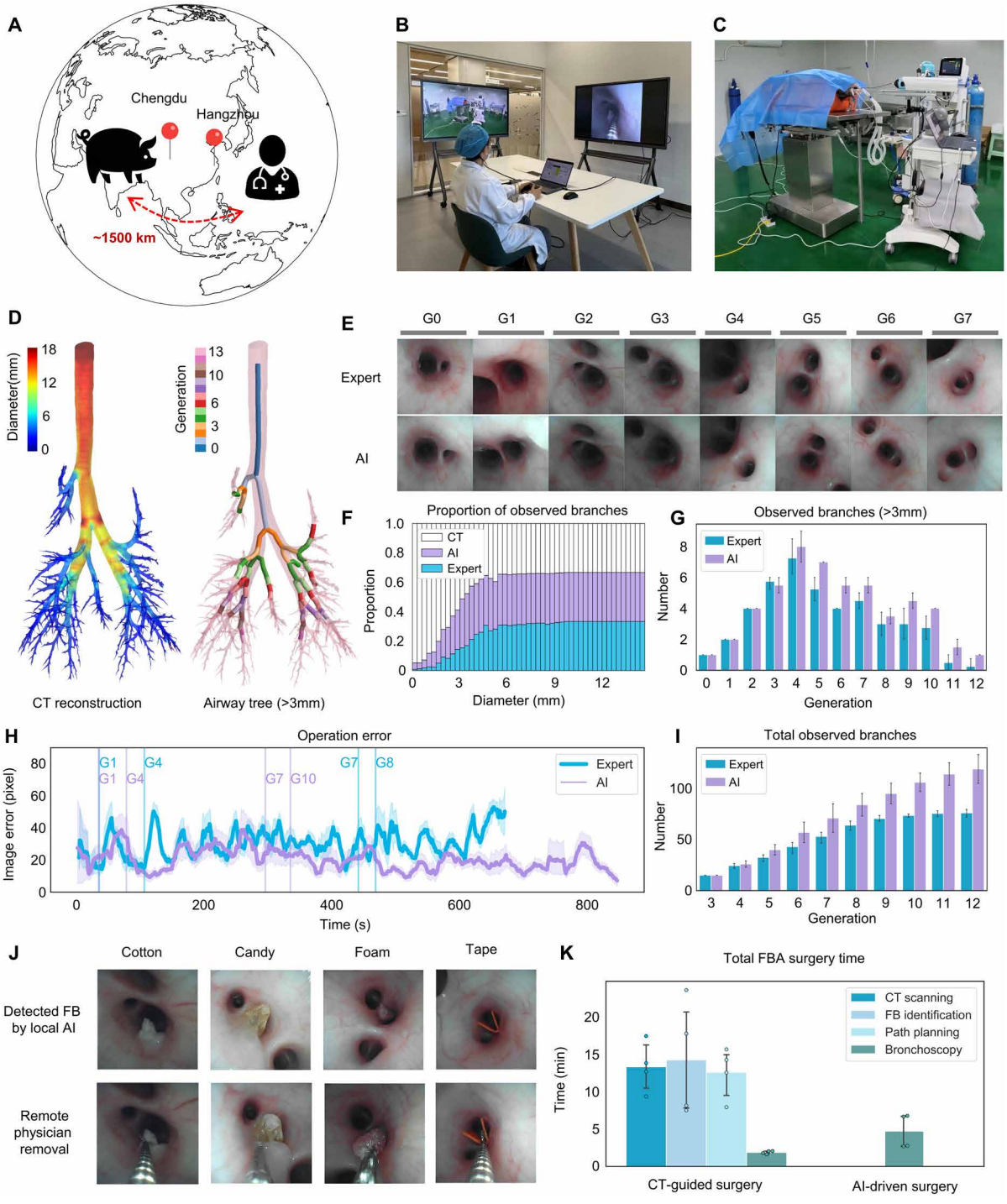


Fig. 5. Results of in vivo experiments. (**A**) Experimental setup. An expert was invited to remotely perform the FB removal procedure from a location 1500 km away from the pig's operating room. (**B**) Remote control scenario for expert operation. (**C**) Local experimental scenario. (**D**) Reconstructed bronchus model from pig CT scan for evaluation only. (**E**) Images captured by the bronchoscope. (**F**) Comparison of three airway exploration methods: CT, AI-guided exploration, and expert-guided exploration. (**G**) Number of observed bronchial branches (>3 mm). Bar plots with error bars depict the mean values ± SD, with $n = 4$ independent experiments. AI-guided exploration outperformed experts in observing deeper and more airways. (**H**) Operational error comparison. All image frames were recorded during the exploration process to compute image errors for comparing operation accuracy between the AI and human expert. The numbers of recorded frames were $n = 20{,}009$ for the expert group and $n = 13{,}228$ for the AI group. Line plots with error bands show time-varying mean image errors ± 95% CI within a time window of 10 frames during each exploration process. (**I**) Total number of observed branches. Bar plots with error bars show the mean values ± SD, with $n = 4$ independent experiments. (**J**) Images of detected FBs by local AI and FB removal by remote physician. (**K**) Comparison of total surgery time between CT-guided surgery and AI-driven surgery. For each type of surgery, bar plots with error bars represent the mean time ± SD, with $n = 4$ independent experiments.

The system-level results presented in this paper demonstrate that AI-supported medical robots can surpass human performance in certain tasks. In the FB search subtask, our AI outperformed human operators in both coverage and accuracy, reducing the likelihood of missing an FB and minimizing patient discomfort during the search. We successfully demonstrated the entire FB removal process in live pigs. Given the animals' lower tolerance to anesthesia, we safely conducted three experiments, proving that our AI can adapt to real-world conditions without requiring extensive expert-labeled datasets. These findings advance the field of medical robotics by highlighting the importance of simulating in vivo conditions and incorporating sensory feedback, such as force and touch, into robotic systems—a crucial aspect for future medical applications.

This case study illustrates the potential of combining AI and robotics for subtasks that require high operational skills but limited medical knowledge. Compared with the time to train a skilled bronchoscopy physician, AI only takes mere days with the aid of graphics processing unit parallelization while delivering a search performance beyond what is feasible for human training. The process is both faster and more cost-effective. Considering that emergency care, and much of everyday health care, consists of routine procedures rather than complex ones, the application of AI to automate such tasks is both feasible and valuable. Thus, for many medical applications not now receiving extensive attention, the autonomy of procedural subtasks and the reduction of cognitive load on physicians represent promising directions for future development in AI and medical robotics.

According to the limit testing of the FB removal of our system (see Supplementary Methods for details), when the FB was much larger (≥10 mm) than the gripper's opening (~3 mm), the success rate of grasp was about 70%. If the size of the FB was close (3 to 5 mm) to the gripper's opening, then the success rate increased considerably (>90%). However, if the object was much smaller (1 mm) than the gripper's opening, then the success rate decreased to ~40%. In contrast, the area under the curve (AUC) of FB identification remained high (>90%) with respect to FBs from 17 to 3 mm. When the size of an FB was 1 mm, the AUC slightly decreased to ~80%. Thus, the bottleneck of the performance of our FB removal system lies in the success rate of grasp. For our system, the limit size of an FB was 1 mm. In clinical practice, this tiny FB was unusual, and most reported sizes of FBs were more than 3 mm (*49*–*51*).

Statistically, around 85% of FBs are located in the bronchi (≥generation 1) and around 15% in the trachea (generation 0) (*2*, *52*). For pediatric patients, because of the immature airway anatomy, most FBs (~63%) are found in the main bronchus (generation 1), with a proportion of 20.4% located in the deeper bronchi (generations 2 to 23). For adult patients, FBs are predominantly found (~57.7%) in deeper bronchi (generations 2 to 23). According to the statistical data, the superiority of our AI-driven surgery in deeper bronchi has a notable clinical value for both children and adults with FBA.

Our robotic system is a nonhaptic teleoperation system, where the force sensor is not equipped for haptic feedback. This setting is the same as that of existing commercial robotic bronchoscopy platforms like Ion and Monarch, which is proven to be acceptable in clinical practice of robotic bronchoscopy. For haptic teleoperation (*53*–*55*), the existence of force feedback imposes higher demands on stale connection to avoid the divergency and unsafety of human-in-the-loop control (*56*, *57*). In our nonhaptic teleoperation, only image streams and human command streams are transferred via 5G communication. For this reason, a more relaxed requirement of bandwidth may be acceptable for our system. As a sacrifice, the velocity and flexibility of direct nonhaptic teleoperation may be much lower than haptic teleoperation, which is worth improving in the future.

This work highlights that valuable medical experts may be dedicating too much time to procedural tasks that could be performed more effectively by robots, thereby underutilizing their true medical expertise. This mismatch between operational demands and expert knowledge is not limited to emergency care but extends across the broader health care system, leading to inefficiencies and resource misallocation for experts, patients, and health care providers. By shifting the role from mere assistants to autonomous collaborators, AI-driven robots have the potential to reshape clinical emergency care and reallocate medical resources, optimizing both health care delivery and outcomes.

MATERIALS AND METHODS

Overview of the hardware for the AI-human collaboration system

The hardware system comprised a portable mobile platform and a bronchoscope robot connected via an electric slide mechanism. This mobile platform was customizable to specific procedural needs and was easily maneuverable within the operating room. The electric slide drove the bronchoscope's feed mechanism during surgery. The bronchoscope robot, weighing ~2 kg (figs. S1 and S21B), consisted of a catheter with a high-stiffness transition section (primarily featuring a braided mesh structure) and a low-stiffness steering section (composed of hinge structures). The catheter contained a working channel (inner diameter: 1.2 mm), a miniaturized endoscopic camera, and two LED lights. The steering mechanism used four linear motors, each controlling one of the catheter's tendons, allowing coordinated steering motions (fig. S2, A and B). The bending angle of the catheter was controlled in the tendon-driven method, where each pair of symmetric linear motors independently controlled the pitch angle and the yaw angle, allowing for omnidirectional bending. As shown in fig. S3, we characterized the bending angles of the catheter under different tendon displacements in free space, which reflected the kinematic characteristics. A linear relationship could be observed without any contact and shape change of the catheter. When the biopsy forceps were inserted, the bending performance of the catheter decreased slightly. The forceps mechanism used two additional linear motors: one for advancing the biopsy forceps and another for actuating its opening and closing (outer diameter: 1 mm), enabling the retrieval of FBs (fig. S2, C and D). The modular system design allowed for straightforward disassembly and replacement of the forceps, enhancing the system's flexibility and ease of maintenance. With a total cost of less than 5000 USD (table S2), the bronchoscopy robot was affordable for local clinics, promoting timely and high-quality care for emergency patients.

The AI-human collaboration system consisted of three key components: the bronchoscope robot hardware, as shown in fig. S4A; the local AI agent; and the remote physician. During the procedure, local clinical staff first inserted the bronchoscope into the patient's airway. The AI agent then autonomously controlled the robot for FB search, receiving real-time bronchoscopic images from the robot's camera and sending steering and advancement/retraction actions to the robot via the serial ports. In the search process, bronchoscopic

images were sent to a remote physician for supervision via 5G communication. The physician retained the authority to override the AI agent, intervening when necessary. Upon detecting a suspected FB, the AI agent issued an alert to the physician, who determined the appropriate course of action. If removal was required, the physician directly controlled the bronchoscope (fig. S6B), sending real-time control commands via 5G while continuously receiving real-time imaging feedback.

Policy neural network architecture

The input of the policy neural network consisted of a reference image of bifurcation, a guidance image from TLMB, and a sequence of observed bronchoscopic images. As shown in fig. S12, three feature extractors, where i indicates the index of each feature extractor, $\{\Phi_E^i\}_{i=1}^3$ with shared parameters were designed to process these three types of input images, respectively, extracting three 512D feature vectors. Then, these features were concatenated to form a 1536D feature as the common feature for inputting the following bifurcation head Φ_B and steering head Φ_S. For FB identification, the latest bronchoscopic image was solely encoded by the feature extractor into a 512D feature for the FB head Φ_F, as shown in fig. S18. In this study, the feature extractor was based on ResNet-34, and these three task-specific heads were based on multilayer perceptrons (MLPs). Note that the input channel number of the feature extractor was matched with the concatenated channel of 10 frames of an image. When the feature extractor received a single image as input, the image was automatically repeated to form a 10-frame sequence.

TLMB design

The core of the TLMB consisted of a compact tree-like topological map and a DFS-based planner. The tree-like topological map encoded a compressed topological representation of the bronchial tree, where nodes stored information about explored bifurcations and unexplored frontiers and edges denoted the connectivity between these nodes. As shown in Fig. 2A, gray nodes represent frontiers, whereas yellow nodes indicate bifurcations that have already been explored. The DFS-based planner was responsible for determining the exploration order of frontiers in the tree-like topological map, ensuring an efficient and comprehensive search of the bronchial tree. The overall bronchial exploration process was divided into three phases: tree construction, forward frontier exploration, and backward node retracing.

Tree construction

At the beginning of exploration, an initial node was established that stored a bronchoscope image as the starting reference image. The lumen detection was performed on the reference image to identify the number and position of airway lumens in the 2D image coordinate system. These detected lumens served as the unexplored branches, namely, frontiers. For each detected lumen, a Gaussian-blurred image was generated centered on the detected position, which was considered as the guidance image of the corresponding frontier. Both the guidance images and the initial reference image were stored in the current node. Each frontier was assigned a unique identity (ID) and pushed into a first-in-first-out queue. Subsequently, the DFS-based planner popped a frontier from the queue to explore, according to the DFS algorithm. The policy neural network took as input both the exploration guidance (the guidance image of the frontier and the reference image of the current node) and observations (incoming bronchoscopic images). Then, the network output steering actions to control the robot and simultaneously evaluated whether a new bifurcation had been reached. Upon detecting a bifurcation, a new node was established, and an edge was formed linking this node to the former one.

Forward frontier exploration

In this phase, the robot was situated at an established node, and the DFS-based planner selected a frontier from the queue to explore, providing the exploration for the policy neural network. During robot exploration from this node, all observed bronchoscope images were temporarily stored in a buffer until the next node was established. From this buffer, 10 images were uniformly sampled to constitute the observation input for the policy neural network. The details of the sampling strategy can be found in Supplementary Methods. The network then output steering actions that controlled the robot's turning. Note that the advancement of the robot was determined by the TLMB. In the forward frontier exploration phase, the advancement of the robot was fixed to a positive velocity. To record the distance of robot advancement, the time from one node to the next node was recorded and saved, facilitating the later backward node retracing process. After a frontier was selected, its ID was removed from the queue. The exploration process concluded when the frontier queue was entirely depleted.

Backward node retracing

During exploration, if the depth of the tree-like topological map reached a threshold, then a leaf node was established, triggering a backward retracing operation to previous nodes. The target node to backtrack was determined by the DFS-based planner, which identified the nearest node having unexplored frontiers. Then, the planner planned a path to a target node and computed the total time of backtracking from the current node to the target node according to the stored advancement time in each node. The reference image of the target node and the real-time image from the camera were inputted into a policy neural network for safe steering. The velocity of backtracking was opposite to that of forward exploration. As an additional safety measure, remote experts were permitted to intervene during the exploration phase: If a remote expert deemed that the current depth of the bronchus was sufficient, then they could manually designate the current position as a leaf node, thereby triggering an automatic backward retracing operation.

Note that in the fifth stage of the AI-human collaboration FBA surgery, after the FB was grasped, backward node retracting was also triggered to extract the bronchoscope with the FB out of the bronchus. Because of the existence of the FB in the camera view, the policy network was unstable to control the robot. Alternatively, the tendon displacement values of the robot were stored in the nodes during forward exploration, and these tendon displacement values were queried from nodes and used to autonomously steer the robot during the backward retracting phase (movie S1).

Training strategy

In this work, we adopted an imitation learning framework to train the steering action and bifurcation identification of the policy network. Specifically, an expert policy π^* was designed to supervise the learning of our policy neural network. Executing in the simulation environment, π^* generated a dataset D comprising state-guidance-action triples (s, g, a), where s represented the observed bronchoscopic images, g denoted the exploration guidance, and $a = \pi^*(s, g)$ was the expert's action, denoted as two steering angles and a bifurcation label $(\Delta\theta, \Delta\varphi, y^{\text{bifur}})$. The goal was to train a policy network π,

parameterized by $\boldsymbol{\tau}$, that mapped a given s and g to a predicted action $\hat{a} = \pi(s, g; \boldsymbol{\tau}) = \left(\Delta\hat{\theta}, \Delta\hat{\varphi}, p^{\text{bifur}}\right)$, closely matching the expert action a. By minimizing the loss function Loss_a, the optimal parameters $\boldsymbol{\tau}^*$ were obtained

$$\boldsymbol{\tau}^* = \underset{\boldsymbol{\tau}}{\text{argmin}} \sum_{i}^{N} \text{Loss}_a\left(\pi^*\left(s_i, g_i\right), \pi\left(s_i, g_i; \boldsymbol{\tau}\right)\right) \tag{1}$$

where N is the size of dataset D. The loss function of action was defined as

$$\text{Loss}_a = \text{Loss}_{\text{bifur}} + \gamma \text{Loss}_{\text{steer}} \tag{2}$$

where γ is the balanced parameter between the combination of bifurcation recognition loss and steering action loss. In addition, the FB identification was trained via a supervised learning framework using the FB identification loss Loss_{fb}.

Expert agent

In conventional imitation learning, the expert policy was executed by human experts to collect and annotate training data, a process that was labor intensive and prone to introducing the inconsistency of annotation between different human experts. To address this issue, we designed an expert agent to autonomously execute expert policy in the simulation environment, which followed the same safe operation principle of keeping the bronchoscope centered in the airway by human experts during bronchoscopy. For the expert agent, the centerlines of the bronchial tree were extracted as the reference paths for data annotation. The expert agent autonomously provided the exploration guidance g and the ground-truth action a for state s, substantially reducing the burden on human demonstration. We implemented the dataset aggregation algorithm (DAgger) for imitation learning. Specifically, the initial dataset was constructed by the expert agent to control the simulated bronchoscope robot to follow the airway centerline, with the expert agent generating the ground-truth actions and exploration guidance. The dataset was then expanded through on-policy training, where the policy neural network π had a probability of p (0.5 in this paper) to control the simulated robot to collect data in the simulation environment, and the expert agent annotated these data frame by frame. During the process, the policy network was trained by the Loss_a, which was the combination of bifurcation recognition loss and steering action loss.

Bifurcation recognition loss

In the simulation environment, as shown in Fig. 2B, we assumed that the distance was d from the bifurcation point to the nearest point of the robot on the 3D centerline. If d was less than a threshold d_{bifur}, then the expert agent annotated the current frame as a bifurcation image and projected the 3D bifurcation points of next generation onto the current 2D image as the exploration guidance for the policy network. Thus, the bifurcation recognition loss could be calculated as

$$\text{Loss}_{\text{bifur}} = -\frac{1}{N} \sum_{i=1}^{N} y_i^{\text{bifur}} \cdot \log\left(p_i^{\text{bifur}}\right) \tag{3}$$

where y_i^{bifur} represents the ground-truth label of bifurcation and p_i^{bifur} is the probability of bifurcation predicted by the policy neural network.

Steering action loss

As shown in fig. S13, the expert agent generated the ground-truth steering action using the privileged information including the robot's relative position and orientation to the reference path, namely, the airway centerline. At the current time t, the ground-truth steering action $a_s^t = (\Delta\theta, \Delta\varphi)$ was calculated as follows

$$\Delta\theta = \arccos\left(\frac{\overrightarrow{\mathbf{O}_{\text{cam}}\mathbf{P}_{\text{target}}} \cdot \boldsymbol{z}}{\overline{\mathbf{O}_{\text{cam}}\mathbf{P}_{\text{proj}}}}\right) \tag{4}$$

$$\Delta\varphi = \arcsin\left(\frac{\overline{\mathbf{P}_{\text{target}}\mathbf{P}_{\text{proj}}}}{\overline{\mathbf{O}_{\text{cam}}\mathbf{P}_{\text{target}}}}\right) \tag{5}$$

where $\mathbf{P}_{\text{target}}$ represents the target waypoint on the centerline that the robot should head toward in the next step, $\mathbf{O}_{\text{cam}}$ is the origin point of the camera's coordinate system, and $\mathbf{P}_{\text{proj}}$ is the projection point of $\mathbf{P}_{\text{target}}$ onto the $\text{xO}_{\text{cam}}\mathbf{y}$ plane. $\mathbf{z}$ represents the direction vector along the z axis of the camera's coordinate system. The position of $\mathbf{P}_{\text{target}}$ can be determined on the basis of the current position of the robot and a predefined distance d_{target} along the centerline. Specifically, the nearest waypoint $\mathbf{P}_{\text{nearest}}$ on the centerline relative to the robot's head was identified. Subsequently, the $\text{index}_{\mathbf{P}_{\text{target}}}$, which corresponds to the position of $\mathbf{P}_{\text{target}}$ among all waypoints on the centerline, was computed using the following formula

$$\text{index}_{\mathbf{P}_{\text{target}}} = \underset{m}{\text{argmin}} \left| \sum_{k=\text{index}_{\text{nearest}}}^{m} \overline{\mathbf{P}_k\mathbf{P}_{k+1}} - d_{\text{target}} \right| \tag{6}$$

where $\mathbf{P}_k$ refers to a specific waypoint on the centerline and $\overline{\mathbf{P}_k\mathbf{P}_{k+1}}$ represents the Euclidean distance between $\mathbf{P}_k$ and its adjacent waypoint $\mathbf{P}_{k+1}$.

On the basis of this, considering the predicted steering action $\hat{a}_s^t = \left(\Delta\hat{\theta}, \Delta\hat{\varphi}\right)$, the steering action loss can be formulated as a mean square error loss as

$$\text{Loss}_{\text{steer}} = \frac{1}{N} \sum_{i=1}^{N} \|a_s^i - \hat{a}_s^i\|_2^2 \tag{7}$$

FB identification loss

Because of the imbalanced distribution of FB and nonobject images during exploration, we implemented a balanced training strategy for the foreign object detection head. We randomly placed various objects, such as cotton and toys, inside the simulation environment and then performed balanced sampling to collect a comparable number of FB and nonobject images, forming an offline dataset. These images were annotated for supervised training of the foreign object identification network. For the first iteration, the policy neural network was trained using Loss_a. Then, the network parameters of the trained feature extractor were shared by another feature extractor that was connected to the FB identification head based on MLP. For the next iteration, the network was trained on the offline dataset using Loss_{fb}. The process was repeated until convergence was achieved. The FB identification loss was defined as

$$\text{Loss}_{\text{fb}} = -\frac{1}{N} \sum\nolimits_{i=1}^{N} y_i^{\text{fb}} \cdot \log\left(p_i^{\text{fb}}\right) \tag{8}$$

where y_i^{fb} represents the ground-truth label of the FB and p_i^{bifur} is the probability of bifurcation predicted by the network.

Sim2Real adaptation and domain randomization

To enhance the effectiveness of the policy neural network in clinical settings, it was essential to apply domain adaptation techniques that reduced discrepancies between simulated and real environments. During the training phase, we performed the image translation on the input images of the policy network, imposing a more realistic image style to the network for Sim2Real adaptation. The details of Sim2Real adaptation are described in Supplementary Methods. To improve the generalization of the policy neural network, two domain randomization methods were designed in this study. To simulate the coupling of respiration with bronchoscope motion, periodic motions were introduced into the simulation environment. The frequency and amplitude of these motions were randomly sampled within a range to simulate both stable and rapid breathing scenarios, enhancing the policy network's generalization. In addition, simulated FBs were randomly placed at different positions within the airway, with variations in type, texture, and appearance. Similarly, data augmentation was also performed on image color, brightness, contrast, and saturation.

Animal studies

We used one female miniature Bama pig, 5 months old, with a body weight of ~38.5 kg at the time of the experiment. The pig was 1.3 m long, in good health, and had a complete vaccination history, including vaccinations against classical swine fever, pseudorabies, porcine reproductive and respiratory syndrome, mycoplasma pneumonia, and foot-and-mouth disease. Before the experiment, the animal study was approved by Chengdu Dossy Experimental Animal Co. Ltd. under license no. 51203500055319, conforming to the National Institutes of Health Guide for the Care and Use of Laboratory Animals. The animal was used to evaluate the performance of the AI-assisted bronchoscopy robot designed for airway exploration and foreign object removal. The robot's control accuracy and operational safety were tested in vivo under physiologically relevant conditions.

At the beginning of the surgical procedure, general anesthesia was induced and maintained using a combination of intravenous and inhalation methods. The animal was intubated and connected to a mechanical ventilator to ensure adequate oxygenation and anesthetic depth throughout the operation. Anesthesia depth was continuously monitored, and additional anesthetic agents were administered as needed. This protocol was consistently applied across all experimental sessions.

Postoperative care included bronchoscopic examinations followed by administration of cefazolin injections to prevent infection. The animal was not euthanized after the experiment and was returned to normal husbandry conditions, where it remains under routine care.

Statistical analysis

Statistical analysis was conducted using Python (version 3.8), NumPy (version 1.24.1), and MATLAB (R2024a). Quantitative results were obtained by comparing AI-guided and expert-guided bronchoscopy procedures across multiple outcome metrics, including centerline deviation, number of branches explored, and time to reach the target FBs. P values were reported, and differences were considered statistically significant if $P < 0.05$.

Supplementary Materials

The PDF file includes:

Supplementary Methods
Figs. S1 to S23
Tables S1 to S3
Reference (*58*)

Other Supplementary Material for this manuscript includes the following:

Movies S1 to S8
MDAR Reproducibility Checklist

REFERENCES AND NOTES

1. A. Jaswal, U. Jana, P. K. Maiti, Tracheo-bronchial foreign bodies: A retrospective study and review of literature. *Indian J. Otolaryngol. Head Neck Surg.* **66**, 156–160 (2014).
2. A. B. Ulas, Y. Aydin, A. Eroglu, Foreign body aspirations in children and adults. *Am. J. Surg.* **224**, 1168–1173 (2022).
3. S. Ngamsanga, V. Vathanophas, K. Ungkanont, A. Tanphaichitr, T. Wannarong, Pediatric respiratory tract foreign bodies in children: A systematic review. *Auris Nasus Larynx* **50**, 607–613 (2023).
4. G. Mosnaim, Asthma in adults. *N. Engl. J. Med.* **389**, 1023–1031 (2023).
5. J. L. Antón-Pacheco, R. Martín-Alelú, M. López, R. Morante, L. Merino-Mateo, S. Barrero, R. Castilla, I. Cano, A. García, A. Gómez, M. C. Luna-Paredes, Foreign body aspiration in children: Treatment timing and related complications. *Int. J. Pediatr. Otorhinolaryngol.* **144**, 110690 (2021).
6. X. Zhou, L. Wang, Y. Zhang, Q. Wu, Y. Cao, The Heimlich maneuver and chest compression relieve mask ventilation difficulties caused by asymptomatic laryngeal mass: A case report and review of literature. *Medicine* **102**, e36362 (2023).
7. S. Martínez-Isasi, A. Carballo-Fazanes, C. Jorge-Soto, M. Otero-Agra, F. Fernández-Méndez, R. Barcala-Furelos, V. Izquierdo, M. García-Martínez, A. Rodríguez-Núñez, School children brief training to save foreign body airway obstruction. *Eur. J. Pediatr.* **182**, 5483–5491 (2023).
8. C. Freitas, M. Serino, C. Cardoso, S. Saleiro, A. P. Vaz, H. Novais-Bastos, A. Morais, A. Magalhães, G. Fernandes, Predictors of survival and technical success of bronchoscopic interventions in malignant airway obstruction. *J. Thorac. Dis.* **13**, 6760–6768 (2021).
9. A. T. N. Ho, A. Shah, A. E. S. Sagar, Review of the clinical outcomes of therapeutic bronchoscopy for central airway obstruction. *Mediastinum* **7**, 17 (2023).
10. C. Zhang, D. M. Dibardino, K. C. Ma, Bronchoscopic treatment of thoracic malignancy. *Med. J.* **8**, 10.21037/amj-23-102 (2023).
11. J. P. Wachs, A. W. Kirkpatrick, S. A. Tisherman, Procedural telementoring in rural, underdeveloped, and austere settings: Origins, present challenges, and future perspectives. *Annu. Rev. Biomed. Eng.* **23**, 115–139 (2021).
12. A. Al-Ibraheem, A. S. Abdlkadir, A. Mohamedkhair, M. Mikhail-Lette, M. Al-Qudah, D. Paez, A. H. Mansour, Cancer diagnosis in areas of conflict. *Front. Oncol.* **12**, 1087476 (2022).
13. A. Shademan, R. S. Decker, J. D. Opfermann, S. Leonard, A. Krieger, P. C. W. Kim, Supervised autonomous robotic soft tissue surgery. *Sci. Transl. Med.* **8**, 337ra64 (2016).
14. A. I. Chen, M. L. Balter, T. J. Maguire, M. L. Yarmush, Deep learning robotic guidance for autonomous vascular access. *Nat. Mach. Intell.* **2**, 104–115 (2020).
15. H. Saeidi, J. D. Opfermann, M. Kam, S. Wei, S. Leonard, M. H. Hsieh, J. U. Kang, A. Krieger, Autonomous robotic laparoscopic surgery for intestinal anastomosis. *Sci. Robot.* **7**, eabj2908 (2022).
16. P. E. Dupont, B. J. Nelson, M. Goldfarb, B. Hannaford, A. Menciassi, M. K. O'Malley, N. Simaan, P. Valdastri, G. Z. Yang, A decade retrospective of medical robotics research from 2010 to 2020. *Sci. Robot.* **6**, eabi8017 (2021).
17. C. D'Ettorre, A. Mariani, A. Stilli, F. Rodriguez y Baena, P. Valdastri, A. Deguet, P. Kazanzides, R. H. Taylor, G. S. Fischer, S. P. Di Maio, A. Menciassi, D. Stoyanov, Accelerating surgical robotics research: A review of 10 years with the da Vinci Research Kit. *IEEE Robot. Autom. Mag.* **28**, 56–78 (2021).
18. M. F. Keating, J. Zhang, C. L. Feider, S. Retailleau, R. Reid, A. Antaris, B. Hart, G. Tan, T. E. Milner, K. Miller, L. S. Eberlin, Integrating the MasSpec pen to the da Vinci Surgical System for in vivo tissue analysis during a robotic-assisted porcine surgery. *Anal. Chem.* **92**, 11535–11542 (2020).
19. M. Roche, The MAKO robotic-arm knee arthroplasty system. *Arch. Orthop. Trauma Surg.* **141**, 2043–2047 (2021).
20. J. Zhang, N. Ng, C. E. H. Scott, M. J. G. Blyth, F. S. Haddad, G. J. Macpherson, J. T. Patton, N. D. Clement, Robotic arm-assisted versus manual unicompartmental knee arthroplasty : A systematic review and meta-analysis of the MAKO robotic system. *Bone Jt. J.* **104-B**, 541–548 (2022).
21. S. M. P. Rossi, F. Benazzo, Individualized alignment and ligament balancing technique with the ROSA® robotic system for total knee arthroplasty. *Int. Orthop.* **47**, 755–762 (2023).

22. D. Fontaine, F. Almairac, B. Chiapello, A. Leplus, Application accuracy of Neuromate robot-guided deep brain stimulation procedures using the non-invasive frameless Neurolocate registration system. *Int. J. Med. Robot.* **20**, e2610 (2024).
23. A. Kuntz, M. Emerson, T. E. Ertop, I. Fried, M. Fu, J. Hoelscher, M. Rox, J. Akulian, E. A. Gillaspie, Y. Z. Lee, F. Maldonado, R. J. Webster III, R. Alterovitz, Autonomous medical needle steering in vivo. *Sci. Robot.* **8**, eadf7614 (2023).
24. M. Johnson, C. Sims, "Bronchoscopy and removal of foreign bodies from the trachea" in *A Guide to Pediatric Anesthesia*, C. Sims, D. Weber, C. Johnson, Eds. (Springer, 2020), pp. 351–363.
25. S. Tanaka, K. Yoshida, K. Muramatsu, S. Yamagishi, S. Obara, K. Watanabe, An experience of subglottic airway foreign body removal in a patient under tracheal intubation: A case report. *JA Clin. Rep.* **6**, 76 (2020).
26. J. Han, J. Davids, H. Ashrafian, A. Darzi, D. S. Elson, M. Sodergren, A systematic review of robotic surgery: From supervised paradigms to fully autonomous robotic approaches. *Int. J. Med. Robot.* **18**, e2358 (2022).
27. A. A. Gumbs, I. Frigerio, G. Spolverato, R. Croner, A. Illanes, E. Chouillard, E. Elyan, Artificial intelligence surgery: How do we get to autonomous actions in surgery? *Sensors* **21**, 5526 (2021).
28. J. Reisenauer, M. J. Simoff, M. A. Pritchett, D. E. Ost, A. Majid, C. Keyes, R. F. Casal, M. S. Parikh, J. Diaz-Mendoza, S. Fernandez-Bussy, E. E. Folch, Ion: Technology and techniques for shape-sensing robotic-assisted bronchoscopy. *Ann. Thorac. Surg.* **113**, 308–315 (2022).
29. M. J. Simoff, M. A. Pritchett, J. S. Reisenauer, D. E. Ost, A. Majid, C. Keyes, R. F. Casal, M. S. Parikh, J. Diaz-Mendoza, S. Fernandez-Bussy, E. E. Folch, Shape-sensing robotic-assisted bronchoscopy for pulmonary nodules: Initial multicenter experience using the Ion™ Endoluminal System. *BMC Pulm. Med.* **21**, 322 (2021).
30. S. K. Iwamoto, W. S. Tsai, Novel approaches utilizing robotic navigational bronchoscopy: A single institution experience. *J. Robot. Surg.* **17**, 1001–1006 (2023).
31. E. Ho, G. Hedstrom, S. Murgu, Robotic bronchoscopy in diagnosing lung cancer—the evidence, tips and tricks: A clinical practice review. *Ann. Transl. Med.* **11**, 359 (2023).
32. M. Simoff, H. Bedi, B. Eperjesiova, "Foreign bodies in the airway: Endoscopic methods" in *Interventions in Pulmonary Medicine*, J. P. Díaz-Jiménez, A. N. Rodríguez, Eds. (Springer, 2023), pp. 685–712.
33. Z. Xu, D. Deng, K. Shimada, Autonomous UAV exploration of dynamic environments via incremental sampling and probabilistic roadmap. *IEEE Robot. Autom. Lett.* **6**, 2729–2736 (2021).
34. R. Bai, H. Guo, W. Y. Yau, L. Xie, Graph-based slam-aware exploration with prior topo-metric information. *IEEE Robot. Autom. Lett.* **9**, 7597–7604 (2024).
35. R. Ma, R. Wang, Y. Zhang, S. Pizer, S. K. McGill, J. Rosenman, J. M. Frahm, RNNSLAM: Reconstructing the 3D colon to visualize missing regions during a colonoscopy. *Med. Image Anal.* **72**, 102100 (2021).
36. J. Shan, Y. Li, T. Xie, H. Wang, ENeRF-SLAM: A dense endoscopic SLAM with neural implicit representation. *IEEE Trans. Med. Robot. Bionics* **6**, 1030–1041 (2024).
37. C. Wang, M. Oda, Y. Hayashi, B. Villard, T. Kitasaka, H. Takabatake, M. Mori, H. Honma, H. Natori, K. Mori, A visual SLAM-based bronchoscope tracking scheme for bronchoscopic navigation. *Int. J. Comput. Assist. Radiol. Surg.* **15**, 1619–1630 (2020).
38. S. N. Lavasani, P. Farnia, E. Najafzadeh, S. Saghatchi, M. Samavati, H. Abtahi, M. Deevband, A. Ahmadian, Bronchoscope motion tracking using centerline-guided Gaussian mixture model in navigated bronchoscopy. *Phys. Med. Biol.* **66**, 025001 (2021).
39. J. Sganga, D. Eng, C. Graetzel, D. Camarillo, "Offsetnet: Deep learning for localization in the lung using rendered images" in *2019 International Conference on Robotics and Automation (ICRA)* (IEEE, 2019), pp. 5046–5052.
40. C. Zhao, M. Shen, L. Sun, G.-Z. Yang, Generative localization with uncertainty estimation through video-CT data for bronchoscopic biopsy. *IEEE Robot. Autom. Lett.* **5**, 258–265 (2019).
41. M. Shen, Y. Gu, N. Liu, G.-Z. Yang, Context-aware depth and pose estimation for bronchoscopic navigation. *IEEE Robot. Autom. Lett.* **4**, 732–739 (2019).
42. Y. Gu, C. Gu, J. Yang, J. Sun, G.-Z. Yang, Vision-kinematics interaction for robotic-assisted bronchoscopy navigation. *IEEE Trans. Med. Imaging* **41**, 3600–3610 (2022).
43. A. Banach, F. King, F. Masaki, H. Tsukada, N. Hata, Visually navigated bronchoscopy using three cycle-consistent generative adversarial network for depth estimation. *Med. Image Anal.* **73**, 102164 (2021).
44. H. Liu, S. Zhang, Y. Yang, L. Sun, Z. Song, S. Xu, A robust pose optimization scheme with spatial geometry awareness for hybrid bronchoscopic navigation. *IEEE Trans. Instrum. Meas.* **73**, 1–10 (2024).
45. Y. Zou, B. Guan, J. Zhao, S. Wang, X. Sun, J. Li, Robotic-assisted automatic orientation and insertion for bronchoscopy based on image guidance. *IEEE Trans. Med. Robot. Bionics* **4**, 588–598 (2022).
46. J. Zhang, L. Liu, P. Xiang, Q. Fang, X. Nie, H. Ma, J. Hu, R. Xiong, Y. Wang, H. Lu, AI co-pilot bronchoscope robot. *Nat. Commun.* **15**, 241 (2024).
47. J. A. Gorden, "Rigid bronchoscopy" in *Principles and Practice of Interventional Pulmonology*, A. Ernst, F. J. F. Herth, Eds. (Springer, 2013), pp. 285–295.
48. D. L. Stahl, K. M. Richard, T. J. Papadimos, Complications of bronchoscopy: A concise synopsis. *Int. J. Crit. Illn. Inj. Sci.* **5**, 189–195 (2015).
49. V. B.-C. Hsue, V. Patel, A coin in the airway. *N. Engl. J. Med.* **390**, e33 (2024).
50. M. Maglione, F. Antonelli, C. Orlando, Iron tablet inhalation: Not a foreign body like the others. *Pediatr. Emerg. Care* **38**, e1696–e1697 (2022).
51. I. Adaletli, S. Kurugoglu, S. Ulus, H. Ozer, M. Elicevik, F. Kantarci, I. Mihmanli, C. Akman, Utilization of low-dose multidetector CT and virtual bronchoscopy in children with suspected foreign body aspiration. *Pediatr. Radiol.* **37**, 33–40 (2007).
52. Z. Huankang, X. Kuanlin, H. Xiaolin, D. Witt, Comparison between tracheal foreign body and bronchial foreign body: A review of 1007 cases. *Int. J. Pediatr. Otorhinolaryngol.* **76**, 1719–1725 (2012).
53. J. Louca, K. Eder, J. Vrublevskis, A. Tzemanaki, Impact of haptic feedback in high latency teleoperation for space applications. *ACM Trans. Hum.-Robot Interact.* **13**, 1–21 (2024).
54. D. G. Black, D. Andjelic, S. E. Salcudean, Evaluation of communication and human response latency for (human) teleoperation. *IEEE Trans. Med. Robot. Bionics* **6**, 53–63 (2024).
55. E. Ivanova, J. Eden, S. Zhu, G. Carboni, A. Yurkewich, E. Burdet, Short time delay does not hinder haptic communication benefits. *IEEE Trans. Haptics* **14**, 322–327 (2021).
56. S. Hirche, M. Buss, Human-oriented control for haptic teleoperation. *Proc. IEEE* **100**, 623–647 (2012).
57. R. V. Patel, S. F. Atashzar, M. Tavakoli, Haptic feedback and force-based teleoperation in surgical robotics. *Proc. IEEE* **110**, 1012–1027 (2022).
58. G. Z. Yang, J. Cambias, K. Cleary, E. Daimler, J. Drake, P. E. Dupont, N. Hata, P. Kazanzides, S. Martel, R. V. Patel, V. J. Santos, R. H. Taylor, Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy. *Sci. Robot.* **2**, eaam8638 (2017).

Acknowledgments: We would like to thank Chengdu Dossy Experimental Animals Co. Ltd. for providing technique support for the animal experiment and the Hangzhou Branch of China Telecom for robust 5G communication technology support and after-sales service. **Funding:** This work was supported by the National Natural Science Foundation of China under grant nos. U24A20128 (R.X.), 62373322 (Y.W.), T2293724 (H.L.), and 62303407 (H.L.); the Zhejiang Provincial Natural Science Foundation of China under grant no. LD25F030001 (Y.W.); and the Xiaomi Foundation (H.L.). **Author contributions:** Y.W., H.L., and R.X proposed the conception of the project; L.L. and J.Z. performed all of the experiments and collected the data with the help of F.W., J.Y., Y.C., and Z.L.; J.H. provided technical support for the clinical experiments; Y.W., H.L., and R.X. provided funding; Y.W., H.L., R.X., L.L., and J.Z. wrote the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Data for evaluating the conclusions are presented at https://doi.org/10.5281/zenodo.14882359. Source code is available at https://doi.org/10.5281/zenodo.15749324 and https://github.com/LiuLiluZJU/AI-Human-Collabration-Bronchoscope-Robot-for-FBA.

Submitted 30 September 2024
Accepted 1 July 2025
Published 30 July 2025
10.1126/scirobotics.adt5338

Science Robotics

AI search, physician removal: Bronchoscopy robot bridges collaboration in foreign body aspiration

Lilu Liu, Jingyu Zhang, Fei Wang, Jiyu Yu, Yuxiang Cui, Zhibin Li, Jian Hu, Rong Xiong, Haojian Lu, and Yue Wang

Sci. Robot. **10** (104), eadt5338. DOI: 10.1126/scirobotics.adt5338

Science Robotics (ISSN 2470-9476) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Robotics* is a registered trademark of AAAS.