Advancing Robot Sensing, Planning, and Adaptability for High-Stakes Medical Automation

Zih-Yun Chiu
University of California San Diego

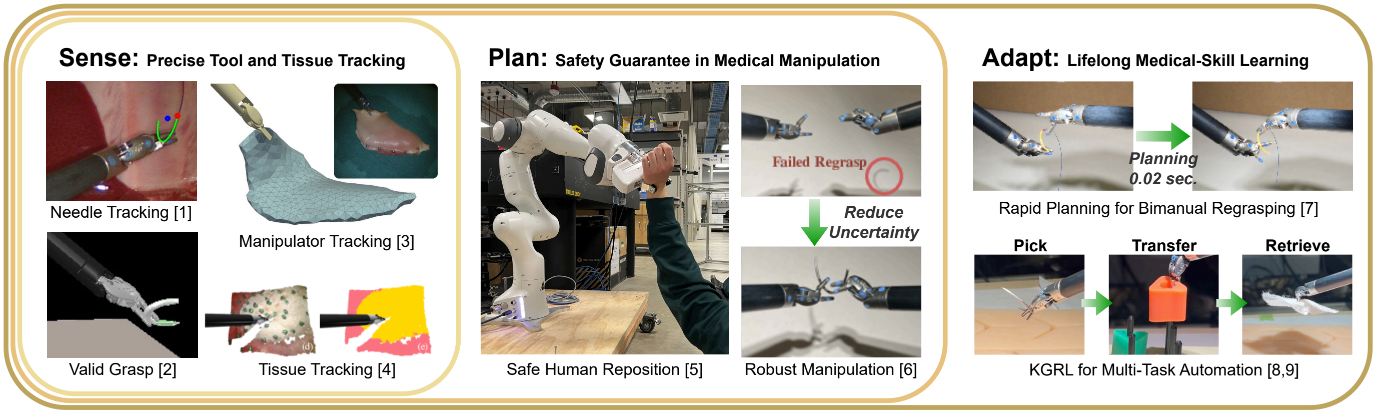

Fig. 1. My past and ongoing research advances robot sensing, planning, and adaptability in medical environments.

I. INTRODUCTION

Advancements in robotics have made precise autonomy in medical procedures possible. Autonomy in robotic medical operations is crucial for delivering treatments to underserved or hazardous areas. My research aims to develop efficient, safe, and generalizable algorithms that enable robots to independently perform safety-critical and time-sensitive medical tasks, thereby democratizing access to quality treatments.

Automating medical tasks such as robotic surgery or human extraction presents several challenges: (1) The robots must be resilient despite environmental uncertainties to achieve the precision and efficiency required for timely interventions. (2) Actions performed on human bodies must be anatomically and biomechanically safe. (3) Robots' strategies should operate across diverse environments, anatomies, and morphologies.

Given these challenges, my research advances robot sensing, planning, and adaptability to achieve autonomous surgical and human-body manipulation (Fig. 1). We developed visual tracking algorithms that enable accurate localization of surgical tools and tissues [1, 2, 3, 4] (sense). We designed safe manipulation methods to ensure human biomechanical safety and minimize surgical visual uncertainty throughout robot operations [5, 6] (plan). We introduced reinforcement learning (RL) frameworks that leverage arbitrary external policies and accumulate knowledge over time, resulting in quick adaptation to new environments and tasks [7, 8, 9] (adapt). Our sensing, planning, and adapting algorithms allow robots to automate over ten medical tasks in real and unstructured environments.

My future research will expand upon the foundation built by our previous studies and explore the following directions:

1) **(Sensing) Situation and Environment Understanding from Multimodal Medical Data**: I aim to leverage multimodal data streams, such as videos, medical images (e.g., CT scans and MRI), haptic feedback, and biosensor data, to make reliable decisions in operations.
2) **(Planning) Advancing Medical Proficiency and Minimizing Risk in Human Manipulation**: I intend to introduce planning algorithms that incorporate systematic medical knowledge and human anatomical and biomechanical models. These algorithms will allow for high-quality procedures and avoid harm to human bodies.
3) **(Adapting) Efficient, Generalizable, and Lifelong Medical-Skill Development**: I seek to enhance the generalizability and *incrementality* of medical agents so that they (1) operate across different conditions and (2) accumulate unstructured knowledge over time to quickly adapt to new environments, tasks, and human bodies.

II. PAST AND ONGOING RESEARCH

Sensing: Uncertainty-, Feasibility-, and Anatomy-Aware Visual Tracking in Surgical Environments. Making safe and effective movements in surgical environments requires precisely localizing critical components, such as surgical tools and various tissues. Prior work employed markers or deep learning models to detect surgical instruments in static images [10, 11, 12, 13, 14, 15] or tracked their motions over time [16, 17, 18, 19, 20]. Another line of research focused on tracking tissue deformation using surface-based assumptions [21, 22, 23, 24, 25]. Building upon these foundations, we integrate sensor uncertainty, object interactions, and semantic information for more robust and accurate surgical localization.

We proposed uncertainty-, feasibility-, and anatomy-aware surgical perception frameworks that achieve unprecedented precision in real-time tool and tissue localization [1, 2, 3, 4]. By incorporating mathematical models of pixel-level uncertainty, convex grasping constraints, and anatomy information into Bayesian filters, our frameworks enable robust surgical tool localization under noisy conditions [1, 3], tool reconstruction ensuring grasping feasibility [2], and deformable tissue tracking that maintains anatomical consistency over time [4]. Our methods achieve sub-millimeter accuracy of surgical manipulator and suture needle localization, can follow slight tremors and shifts in tissues, and enable robots to succeed in automating numerous real-world surgical tasks [7, 6, 9].

Planning: Guaranteeing Safety in Human Manipulation. Despite the enhanced dexterity and manipulability, robots nowadays still lack the systematic knowledge required for medical automation, leaving them relying on teleoperation. Prior work has implemented this knowledge through specialized tools [26, 27, 28, 29] and assistive devices [30, 31, 32, 33, 34, 35]. However, task-specific tools require frequent switching, and devices designed for conscious patients may not translate well to unconscious individuals. Furthermore, while existing approaches to medical automation without specialized devices offer promising directions [13, 36, 37, 38, 15, 39], they usually rely on environment-specific trajectories to account for the inherent uncertainties in medical settings.

We proposed uncertainty-aware mathematical models that describe surgical procedures and human biomechanical structures [7, 6, 5]. These models, derived from long-standing surgical practices and kinesiology literature, quantify the quality of a surgical grasp [7], the environmental visibility from an endoscopic camera [6], and the body reaction forces during robot operations [5]. We integrated these models into RL and constrained optimization frameworks to plan safe and efficient robot trajectories in medical environments. Our real-world experiments achieved (1) regrasping millimeter-scale suture needles with less than 0.1s planning time, (2) increasing automation success by more than fourfold in surgical-tool manipulation, even with a moving endoscopic camera, and (3) safely repositioning human bodies using a robotic manipulator.

Adaptability: Efficient and Lifelong Medical-Skill Learning. Well-trained medical professionals can continually enhance their expertise, staying current with the latest medical advancements. Autonomous robots operating in underserved or hazardous areas should also improve their skills over time to perform medical tasks independently in diverse settings over extended periods. While prior work enables robots to acquire complex medical skills through optimization [28, 40, 41], visual servoing [37, 15, 42], learning from demonstrations [43, 44, 38, 45, 46], and RL [47, 48, 49, 50, 51, 52], they struggle with rapid adaptability for unseen tasks and continuous evolution. The challenge of efficiently accumulating knowledge over a robot's lifetime extends beyond medical automation and remains an open question in general robotics.

We developed RL algorithms that flexibly leverage external guidance to enable efficient and lifelong robot learning. We proposed *the ego-centric state/action spaces* and *the mixed exploration strategy* for RL that harnesses medical knowledge to help learn a generalizable suture-needle manipulation policy [7]. Extending this approach to the broader robot learning domain, we introduced *Knowledge-Grounded RL (KGRL)* and *Knowledge-Inclusive Attention Network (KIAN)* that flexibly accumulate knowledge over tasks and simultaneously achieve sample-efficient, generalizable, compositional, and incremental learning, which are the key foundations for efficient robot learning [53]. Building upon KGRL and KIAN, we then proposed *Surgical Incremental RL (SurgIRL)* [9]. Our SurgIRL framework improves KIAN's learning efficiency and incrementally learns multiple surgical tasks by accumulating medical knowledge over time. We effectively solved ten surgical tasks with SurgIRL and deployed the policies to real-world environments, achieving over 90% success across all tasks.

III. Future Directions

Sensing: Situation and Environment Understanding from Multimodal Medical Data. I aim to broaden robot sensing in medical environments beyond vision to include data across modalities. By pursuing partnerships with clinical partners and search-and-rescue workers, I will first uncover the key data attributes that help understand medical scenarios. Then, I will bring together expertise in medical image, tactile, and biosensor data processing to explore cross-modality feature learning and multimodal neural networks for heterogeneous medical data. This collaboration will provide robots with a comprehensive understanding of medical environments.

Planning: Advancing Medical Proficiency and Minimizing Risk in Human Manipulation. I intend to develop mathematical models that span the broad spectrum of medical operations. First, I will explore how different models, such as finite element, viscoelastic, hyperelastic, fluid dynamics, and position-based dynamics models, can be combined to describe the physical attributes of organs and tissues. Next, I will investigate spatial-temporal modeling methods to transfer medical procedures into robot executable actions. Finally, I will develop algorithms based on generative models, e.g., transformers and stable diffusion, to achieve intention prediction in human-robot collaborative medical scenarios. The models and methods proposed can be naturally embedded into planning, optimization, and learning frameworks to enhance the quality and safety of medical automation.

Adapting: Efficient, Generalizable, and Lifelong Medical-Skill Development. I seek to design robot learning frameworks that integrate the community's efforts for lifelong medical skill learning. Today, the research community is teeming with datasets, robot policies for diverse tasks, and foundation models, offering substantial information on perceiving and navigating the world. Nonetheless, we need algorithms that can function beyond the restrictions set by current models, such as their black-box properties, to harness the underlying information that can be shared across multiple environments and tasks. The future directions in this research line will aim to extract unified knowledge from diverse datasets, policies, and models for continuous knowledge accumulation in learning medical or general robotic skills.

REFERENCES

[1] Zih-Yun Chiu, Albert Z Liao, Florian Richter, Bjorn Johnson, and Michael C Yip. Markerless suture needle 6d pose tracking with robust uncertainty estimation for autonomous minimally invasive robotic surgery. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5286–5292. IEEE, 2022.

[2] Zih-Yun Chiu, Florian Richter, and Michael C Yip. Real-time constrained 6d object-pose tracking of an in-hand suture needle for minimally invasive robotic surgery. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4761–4767. IEEE, 2023.

[3] Christopher D'Ambrosia, Florian Richter, Zih-Yun Chiu, Nikhil Shinde, Fei Liu, Henrik I Christensen, and Michael C Yip. Robust surgical tool tracking with pixel-based probabilities for projected geometric primitives. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 15455–15462. IEEE, 2024.

[4] Shan Lin, Albert J Miao, Jingpei Lu, Shunkai Yu, Zih-Yun Chiu, Florian Richter, and Michael C Yip. Semantic-super: a semantic-aware surgical perception framework for endoscopic tissue identification, reconstruction, and tracking. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4739–4746. IEEE, 2023.

[5] Elizabeth Peiros, Zih-Yun Chiu, Yuheng Zhi, Nikhil Shinde, and Michael C Yip. Finding biomechanically safe trajectories for robot manipulation of the human body in a search and rescue scenario. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 167–173. IEEE, 2023.

[6] Nikhil U Shinde, Zih-Yun Chiu, Florian Richter, Jason Lim, Yuheng Zhi, Sylvia Herbert, and Michael C Yip. Surestep: An uncertainty-aware trajectory optimization framework to enhance visual tool tracking for robust surgical automation. In *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6953–6960. IEEE, 2024.

[7] Zih-Yun Chiu, Florian Richter, Emily K Funk, Ryan K Orosco, and Michael C Yip. Bimanual regrasping for suture needles using reinforcement learning for rapid motion planning. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 7737–7743. IEEE, 2021.

[8] Zih-Yun Chiu, Yi-Lin Tuan, William Yang Wang, and Michael Yip. Flexible attention-based multi-policy fusion for efficient deep reinforcement learning. *Advances in Neural Information Processing Systems*, 36:13590–13612, 2023.

[9] Yun-Jie Ho, Zih-Yun Chiu, Yuheng Zhi, and Michael C Yip. Surgirl: Towards life-long learning for surgical automation by incremental reinforcement learning. *arXiv preprint arXiv:2409.15651*, 2024.

[10] Christophe Doignon, Florent Nageotte, and Michel De Mathelin. Detection of grey regions in color images: application to the segmentation of a surgical instrument in robotized laparoscopy. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)*, volume 4, pages 3394–3399. IEEE, 2004.

[11] Sandrine Voros, Jean-Alexandre Long, and Philippe Cinquin. Automatic detection of instruments in laparoscopic images: A first step towards high-level command of robotic endoscopic holders. *The International Journal of Robotics Research*, 26(11-12):1173–1190, 2007.

[12] Santosh Iyer, Thomas Looi, and James Drake. A single arm, single camera system for automated suturing. In *2013 IEEE international conference on robotics and automation*, pages 239–244. IEEE, 2013.

[13] Claudia D'Ettorre, George Dwyer, Xiaofei Du, François Chadebecq, Francisco Vasconcelos, Elena De Momi, and Danail Stoyanov. Automated pick-up of suturing needles for robotic surgical assistance. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1370–1377. IEEE, 2018.

[14] Yuta Yamazaki, Shingo Kanaji, Takeru Matsuda, Taro Oshikiri, Tetsu Nakamura, Satoshi Suzuki, Yuta Hiasa, Yoshito Otake, Yoshinobu Sato, and Yoshihiro Kakeji. Automated surgical instrument detection from laparoscopic gastrectomy video images using an open source convolutional neural network platform. *Journal of the American College of Surgeons*, 230(5):725–732e1, 2020.

[15] Albert Wilcox, Justin Kerr, Brijen Thananjeyan, Jeffrey Ichnowski, Minho Hwang, Samuel Paradis, Danyal Fer, and Ken Goldberg. Learning to localize, grasp, and hand over unmodified surgical needles. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 9637–9643. IEEE, 2022.

[16] Austin Reiter, Peter K Allen, and Tao Zhao. Feature classification for tracking articulated surgical tools. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012: 15th International Conference, Nice, France, October 1-5, 2012, Proceedings, Part II 15*, pages 592–600. Springer, 2012.

[17] Yusuke Kurose, Young Min Baek, Yuya Kamei, Shinichi Tanaka, Kanako Harada, Shigeo Sora, Akio Morita, Naohiko Sugita, and Mamoru Mitsuishi. Preliminary study of needle tracking in a microsurgical robotic system for automated operations. In *2013 13th international conference on control, automation and systems (ICCAS 2013)*, pages 627–630. IEEE, 2013.

[18] M Ferro, GA Fontanelli, F Ficuciello, B Siciliano, M Vendittelli, et al. Vision-based suturing needle tracking with extended kalman filter. In *Computer/Robot Assisted Surgery workshop*, 2017.

[19] Orhan Özgüner, Ran Hao, Russell C Jackson, Tom Shkurti, Wyatt Newman, and M Cenk Cavusoglu. Three-dimensional surgical needle localization and tracking using stereo endoscopic image streams. In *2018 IEEE international conference on robotics and automation (ICRA)*, pages 6617–6624. IEEE, 2018.

[20] Florian Richter, Jingpei Lu, Ryan K Orosco, and Michael C Yip. Robotic tool tracking under partially visible kinematic chain: A unified approach. *IEEE Transactions on Robotics*, 38(3):1653–1670, 2021.

[21] Oscar G Grasa, Ernesto Bernal, Santiago Casado, Ismael Gil, and JMM Montiel. Visual slam for handheld monocular endoscope. *IEEE transactions on medical imaging*, 33(1):135–146, 2013.

[22] Kristen L Lurie, Roland Angst, Dimitar V Zlatev, Joseph C Liao, and Audrey K Ellerbee Bowden. 3d reconstruction of cystoscopy videos for comprehensive bladder records. *Biomedical optics express*, 8(4):2106–2123, 2017.

[23] Andres Marmol, Artur Banach, and Thierry Peynot. Dense-arthroslam: Dense intra-articular 3-d reconstruction with robust localization prior for arthroscopy. *IEEE Robotics and Automation Letters*, 4(2):918–925, 2019.

[24] Jose Lamarca, Shaifali Parashar, Adrien Bartoli, and JMM Montiel. Defslam: Tracking and mapping of deforming scenes from monocular sequences. *IEEE Transactions on robotics*, 37(1):291–303, 2020.

[25] Juan J Gómez-Rodríguez, José Lamarca, Javier Morlana, Juan D Tardós, and José MM Montiel. Sd-defslam: Semi-direct monocular slam for deformable and intracorporeal scenes. In *2021 IEEE international conference on robotics and automation (ICRA)*, pages 5170–5177. IEEE, 2021.

[26] Russell C Jackson and M Cenk Çavuşoğlu. Needle path planning for autonomous robotic surgical suturing. In *2013 IEEE International Conference on Robotics and Automation*, pages 1669–1675. IEEE, 2013.

[27] Simon Leonard, Kyle L Wu, Yonjae Kim, Axel Krieger, and Peter CW Kim. Smart tissue anastomosis robot (star): A vision-guided robotics system for laparoscopic suturing. *IEEE Transactions on Biomedical Engineering*, 61(4):1305–1317, 2014.

[28] Siddarth Sen, Animesh Garg, David V Gealy, Stephen McKinley, Yiming Jen, and Ken Goldberg. Automating multi-throw multilateral surgical suturing with a mechanical needle guide and sequential convex optimization. In *2016 IEEE international conference on robotics and automation (ICRA)*, pages 4178–4185. IEEE, 2016.

[29] Sahba Aghajani Pedram, Peter Ferguson, Ji Ma, Erik Dutson, and Jacob Rosen. Autonomous suturing via surgical robot: An algorithm for optimal selection of needle diameter, shape, and path. In *2017 IEEE International conference on robotics and automation (ICRA)*, pages 2391–2398. IEEE, 2017.

[30] Carmichael F Ong, Jennifer L Hicks, and Scott L Delp. Simulation-based design for wearable robotic systems: an optimization framework for enhancing a standing long jump. *IEEE Transactions on Biomedical Engineering*, 63 (5):894–903, 2015.

[31] Jesús Ortiz, Tommaso Poliero, Giovanni Cairoli, Eveline Graf, and Darwin G Caldwell. Energy efficiency analysis and design optimization of an actuation system in a soft modular lower limb exoskeleton. *IEEE Robotics and Automation Letters*, 3(1):484–491, 2017.

[32] Lelai Zhou, Yibin Li, and Shaoping Bai. A human-centered design optimization approach for robotic exoskeletons through biomechanical simulation. *Robotics and Autonomous Systems*, 91:337–347, 2017.

[33] Rejin John Varghese, Daniel Freer, Fani Deligianni, Jindong Liu, Guang-Zhong Yang, and R Tong. Wearable robotics for upper-limb rehabilitation and assistance: A review of the state-of-the-art challenges and future research. *Wearable technology in medicine and health care*, pages 23–69, 2018.

[34] Xianlian Zhou. Predictive human-in-the-loop simulations for assistive exoskeletons. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, volume 83983, page V009T09A006. American Society of Mechanical Engineers, 2020.

[35] Lingxing Chen, Chunjie Chen, Zhuo Wang, Xin Ye, Yida Liu, and Xinyu Wu. A novel lightweight wearable soft exosuit for reducing the metabolic rate and muscle fatigue. *Biosensors*, 11(7):215, 2021.

[36] Fangxun Zhong, Yaqing Wang, Zerui Wang, and Yun-Hui Liu. Dual-arm robotic needle insertion with active tissue deformation for autonomous suturing. *IEEE Robotics and Automation Letters*, 4(3):2669–2676, 2019.

[37] Florian Richter, Shihao Shen, Fei Liu, Jingbin Huang, Emily K Funk, Ryan K Orosco, and Michael C Yip. Autonomous robotic suction to clear the surgical field for hemostasis using image-based blood flow detection. *IEEE Robotics and Automation Letters*, 6(2):1383–1390, 2021.

[38] Kim L Schwaner, Iñigo Iturrate, Jakob KH Andersen, Pernille T Jensen, and Thiusius R Savarimuthu. Autonomous bi-manual surgical suturing based on skills learned from demonstration. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4017–4024. IEEE, 2021.

[39] Karthik Dharmarajan, Will Panitch, Muyan Jiang, Kishore Srinivas, Baiyu Shi, Yahav Avigal, Huang Huang, Thomas Low, Danyal Fer, and Ken Goldberg. Automating vascular shunt insertion with the dvrk surgical robot. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6781–6788. IEEE, 2023.

[40] Jingbin Huang, Fei Liu, Florian Richter, and Michael C Yip. Model-predictive control of blood suction for surgical hemostasis using differentiable fluid simulations. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 12380–12386. IEEE, 2021.

[41] Xiao Liang, Fei Liu, Yutong Zhang, Yuelei Li, Shan Lin, and Michael Yip. Real-to-sim deformable object manipulation: Optimizing physics models with residual mappings for robotic surgery. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 15471–15477. IEEE, 2024.

[42] Ki-Hwan Oh, Leonardo Borgioli, Miloš Žefran, Liaohai Chen, and Pier Cristoforo Giulianotti. A framework for automated dissection along tissue boundary. In *2024 10th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, pages 1427–1433. IEEE, 2024.

[43] Jessica J Ji, Sanjay Krishnan, Vatsal Patel, Danyal Fer, and Ken Goldberg. Learning 2d surgical camera motion from demonstrations. In *2018 IEEE 14th International Conference on Automation Science and Engineering (CASE)*, pages 35–42. IEEE, 2018.

[44] Ameya Pore, Eleonora Tagliabue, Marco Piccinelli, Diego Dall'Alba, Alicia Casals, and Paolo Fiorini. Learning from demonstrations for autonomous soft-tissue retraction. In *2021 international symposium on medical robotics (ISMR)*, pages 1–7. IEEE, 2021.

[45] Kento Kawaharazuka, Kei Okada, and Masayuki Inaba. Robotic constrained imitation learning for the peg transfer task in fundamentals of laparoscopic surgery. *arXiv preprint arXiv:2405.03440*, 2024.

[46] Ji Woong Kim, Tony Z Zhao, Samuel Schmidgall, Anton Deguet, Marin Kobilarov, Chelsea Finn, and Axel Krieger. Surgical robot transformer (srt): Imitation learning for surgical tasks. *arXiv preprint arXiv:2407.12998*, 2024.

[47] Ngoc Duy Nguyen, Thanh Nguyen, Saeid Nahavandi, Asim Bhatti, and Glenn Guest. Manipulating soft tissues by deep reinforcement learning for autonomous robotic surgery. In *2019 IEEE International Systems Conference (SysCon)*, pages 1–7. IEEE, 2019.

[48] Vignesh Manoj Varier, Dhruv Kool Rajamani, Nathaniel Goldfarb, Farid Tavakkolmoghaddam, Adnan Munawar, and Gregory S Fischer. Collaborative suturing: A reinforcement learning approach to automate hand-off task in suturing for surgical robots. In *2020 29th IEEE international conference on robot and human interactive communication (RO-MAN)*, pages 1380–1386. IEEE, 2020.

[49] Yafei Ou and Mahdi Tavakoli. Towards safe and efficient reinforcement learning for surgical robots using real-time human supervision and demonstration. In *2023 International Symposium on Medical Robotics (ISMR)*, pages 1–7. IEEE, 2023.

[50] Marco Caianiello, Cristina Iacono, Antonella Imperato, and Fanny Ficuciello. Exploring the use of deep reinforcement learning algorithms for wound-approaching trajectories in robot-assisted minimally invasive surgery. In *2023 21st International Conference on Advanced Robotics (ICAR)*, pages 285–290. IEEE, 2023.

[51] Ke Fan, Ziyang Chen, Giancarlo Ferrigno, and Elena De Momi. Learn from safe experience: Safe reinforcement learning for task automation of surgical robot. *IEEE Transactions on Artificial Intelligence*, 2024.

[52] Mustafa Haiderbhai, Radian Gondokaryono, Andrew Wu, and Lueder A Kahrs. Sim2real rope cutting with a surgical robot using vision-based reinforcement learning. *IEEE Transactions on Automation Science and Engineering*, 2024.

[53] Leslie Pack Kaelbling. The foundation of efficient robot learning. *Science*, 369(6506):915–916, 2020.